

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 15:20:44 ; Search time 3927.32 Seconds
(without alignments)
5616.301 Million cell updates/sec

Title: US-09-488-265-25
Perfect score: 1426
Sequence: 1 tatatgaattcatggcggtg.....ttcgcttaagaattcatata 1426

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_cm.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: gb_ba1.*
17: gb_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vl1.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rol.*
95: gb_rol2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	% Match	Query Length	DB ID	Description	
1	1342.8	94.2	1426	9	AX021809	Sequence
2	790.8	55.5	1350	56	AF295325	Synthetic
3	622.2	43.6	1404	9	A19452	Phytase cDN
4	622.2	43.6	1404	10	I13430	Sequence cDN
5	622.2	43.6	1404	10	I33881	Sequence 33
6	619.6	43.5	1571	9	AX000634	Sequence 19
7	619.6	43.5	1812	15	SFUS9804	Aspergillus
8	618.2	43.4	2000	13	ANPHYAG	216414 A.niger phy

Sat Oct 27 15:25:36 2001

9	618.2	43.4	2363	9	AR018076	AR018076 Sequence	QY	61	catcgggtaccgcttgggtctcgtgtaattctcactcttctgacactgttgacgggtg	120
10	618.2	43.4	2363	9	AR051916	AR051916 Sequence	QY	61	CATCCGGTACCOCCTTGGGTCTCTGTTGTAATTCCTACTCTTGTGACACTGTTGACGGTG	120
11	618.2	43.4	2379	9	AR053934	AR053934 Sequence	Db	61	CATCCGGTACCOCCTTGGGTCTCTGTTGTAATTCCTACTCTTGTGACACTGTTGACGGTG	120
12	618.2	43.4	2379	13	ASNPHTAS	L02421 Aspergillus	QY	121	gttaccatgtttcccaaaaattttctcacttgggtggtgcaatactctccattctctt	180
13	618.2	43.4	2665	13	ASNPHTASE	M94550 Aspergillus	QY	121	gttaccatgtttcccaaaaattttctcacttgggtggtgcaatactctccattctctt	180
14	618.2	43.4	6756	9	AI19451	A19451 phytase gen	Db	121	gttaccatgtttcccaaaaattttctcacttgggtggtgcaatactctccattctctt	180
15	618.2	43.4	6756	10	II13429	II13429 Sequence 31	QY	181	tggtgacgaatctgctattttctcagacggttcccaagggtgtgtagagttactcttc	240
16	613.4	43.0	1528	13	AF218813	AF218813 Aspergill	QY	181	tggtgacgaatctgctattttctcagacggttcccaagggtgtgtagagttactcttc	240
17	599.6	42.0	1515	12	AB022700	AB022700 Aspergill	QY	181	tggtgacgaatctgctattttctcagacggttcccaagggtgtgtagagttactcttc	240
18	597.6	41.9	1590	14	AY013315	AY013315 Aspergill	Db	181	tggtgacgaatctgctattttctcagacggttcccaagggtgtgtagagttactcttc	240
19	595.8	41.8	1553	12	AB042805	AB042805 Aspergill	QY	241	aaatttttcttagacacggtgcttagatatacccaacttcttctaagtctaaagaagtactctg	300
20	580.4	40.7	1931	9	AX000630	AX000630 Sequence	QY	241	aaatttttcttagacacggtgcttagatatacccaacttcttctaagtctaaagaagtactctg	300
21	580.4	40.7	1931	14	ENU59803	U59803 Emericella	Db	241	aaatttttcttagacacggtgcttagatatacccaacttcttctaagtctaaagaagtactctg	300
22	572.4	40.1	1912	9	A46793	A46793 Sequence 11	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
23	549	38.5	1845	15	TTU59802	TTU59802 Talaromyces	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
24	549	38.5	1845	15	AX000632	AX000632 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
25	533.4	37.4	1567	9	AX000636	U60412 Aspergillus	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
26	533.4	37.4	1770	14	ATU60412	A46783 Sequence 1	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
27	533.4	37.4	2327	9	A46783	U59805 Aspergillus	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
28	533.4	37.4	2327	14	ATU59805	U59805 Aspergillus	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
29	512	35.9	1922	10	AX085191	AX085191 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
30	510	35.8	1464	10	AX085193	AX085193 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
31	510	35.8	1584	10	AX085192	AX085192 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
32	412.4	28.9	4898	10	AX085207	AX085207 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
33	356.6	25.0	2200	9	AR031151	AR031151 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
34	292.6	20.5	853	10	AX085213	AX085213 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
35	280.6	19.7	860	10	AX085208	AX085208 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
36	265.2	18.6	3995	9	A46785	A46785 Sequence 3	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
37	265.2	18.6	3995	14	MTU59806	U59806 Myceliophth	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
38	134.4	9.4	950	10	AX085210	AX085210 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
39	90.2	6.3	1117	14	MGR299239	AJ299239 Magnapor	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
40	69.6	4.9	1404	14	SCU19789	U019789 Saccharomyc	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
41	64.8	4.5	56097	15	YSCU9177	U00029 Saccharomyc	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
42	64	4.5	1404	14	SCPH05A	AX01079 Yeast gene	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
43	64	4.5	1904	10	AX072908	AX072908 Sequence	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
44	64	4.5	2934	15	SCYBK093C	Z35962 S.cerevisia	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360
45	64	4.5	3913	14	SCPH035	X01080 Yeast genes	QY	301	ctttgattgaagctatttcaaaaagaacgctactgctttcaagggttaagtagcgtttcttga	360

ALIGNMENTS

RESULT	1	AX021809	1426 bp	DNA	PAT	07-SEP-2000
LOCUS		Sequence 3	from Patent EP0897985.			
DEFINITION		AX021809				
ACCESSION		AX021809.1	GI:10045052			
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
BASE COUNT		338 a	311 c	308 g	469 t	
ORIGIN						
Query Match		94.2%	Score 1342.8;	DB 9;	Length 1426;	
Best Local Similarity		96.4%	Pred. No. 0;			
Matches 1374;		Conservative				
QY	1	tatatgaattcatggcggttcgtcgctactgtccattgccacattgttcggttcca	60			
Db	1	TATATGAATTCATGGCGGTCGTCGCTACTCTCCATTGCCACCTTGTTCGTTCCA	60			


```

QY 1141 gtactaagccattgtctactactcttctgtgaattattgaatactgacggttaccgtg 1200
Db 1141 GTACTGCTCCATTCTTCTACTCTTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTG 1200
QY 1201 cttcttgactgttccattcgtctgtagagcttaccgttgaattgaatgacggttgaagctg 1260
Db 1201 CTTCTTGGAGTGTCCATTCCGGTCTAGAGCTTACGTTGAATGATGCAATGTCAAGCTG 1260
QY 1261 aaaaagaaccattggttagagtttttggtttaacgacagagttgttccattgacggttgg 1320
Db 1261 AAAAGAACCATTTGGTTAGAGTTTGGTTAAGCAGACAGAGTTGTTCCATTGACGGTTG 1320
QY 1321 gtcttgacaagttgggttagatgtaagagagacgacttctgaagtttgaagtttgcgtta 1380
Db 1321 CTGTTGACAAGTTGGGTAGATGATAGAGAGACGACTTCGTTGAAGGTTGCTTTTCGCTA 1380
QY 1381 gactgtggttaactgggaagaattgttgcgttaagaattcatata 1426
Db 1381 GATCTGCTGTAAGTGGCTGAATGTTTCGCTTAAGAATTTCATATA 1426

RESULT 2
AF295325 LOCUS AF295325 1350 bp mRNA SYN 18-OCT-2000
DEFINITION Synthetic construct phytase mRNA, complete cds.
ACCESSION AF295325
VERSION AF295325.1 GI:10732782
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Yang,L., Chen,Z., Bei,J., Liao,L. and Wang,X.
TITLE Synthetic sequence of phytase gene for expression in Pichia pastoris
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1350)
AUTHORS Chen,Z.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-2000) Animal Science Institute, Guangzhou Academy of Agricultural Sciences, Guangzhou, Guangdong 510640, China
REFERENCE 3 (bases 1 to 1350)
AUTHORS Yang,L., Bei,J., Liao,L. and Wang,X.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-2000) Biopharmaceutical Center, Zhongshan University, Guangzhou, Guangdong 510275, China
FEATURES
source
1..1350
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Aspergillus niger sequence modified for expression in Pichia pastoris"
1..1350
/codon_start=1
/transl_table=1
/product="phytase"
/protein_id="AAG22536.1"
/db_xref="GI:10732783"
/translation="MLAVPASRNOSGDTVDGQYOCFSETSHLWQYAPFSLANESV
ISPEVPACGRVTFAQLSRHRCARYPTDSKGYKYSALIELOQNAITFDGKYAFKLTYN
YSLGADLLTPGEOLVNSGKFORVESLITRNIVPIRSSGSRVIAKSKFIEGQF
STKLKDRAPGOSPKIDVIVSEASSNLTDPGCTVEFDELADTVEANFTAFV
PSIKRLENDLSGVTITDEVTYIMDMCSFDTISTSTVDTKLSPFCDLTHDEWIND
YQSLKKYGHGAGNLPUGTQGVGYAMELILARLTHSPVDDTSSNHTLDSPPATPLN
STLFDPSHDNGIISILFALGLYNGKPLSTTVENITOTDGPSSAWTVPPASRLIVE
MMQCAOEPLVRLVNDVRVPLHGGPVDALGRCTRDSFVRGLSFARSGDWAECFA"
BASE COUNT 319 a 354 c 282 g 395 t
ORIGIN

Query Match 55.5%; Score 790.8; DB 56; Length 1350;
Best Local Similarity 74.7%; Pred. No. 2.4e-199;
Matches 993; Conservative 0; Mismatches 337; Indels 0; Gaps 0;

```

```

QY 86 tggtaattctcactcttgtgacactgttgacggtgttaccgaatttccccagaatttc 145
Db 21 TAGAAATCAATCCTCTTGTGATCTGTCGATCAGGGTTATCAATGTTTCTCCGAGACTTC 80
QY 146 tcaacttgggggttcaactactctccattctctcttcttggctgacgaattctgtattctcc 205
Db 81 TCATCTTTGGGGTCAATACGCACCAATCTTCTCTCTGCAACCAATCTGTCACTCTCCCC 140
QY 206 agaagttccaaaggggttagagttacttcttctcaagtttggctagacacggtgctag 265
Db 141 TGAGGTGCCACCGGATGTAGAGTCATTTGCTCAGGTCTCTTCCAGACATGGAGTAG 200
QY 266 ataccacactctcttaagtttaagaagttactctgttctgtatgaaagctattcaagaag 325
Db 201 ATATCCAAACCGACTCCCAAGGTAAGAAATACTCCGCTCTTATTGAGGAGATCCAGAGAA 260
QY 326 cgtactgtcttcaagggtaagtacgcttctcttgaagaattacaacttcttgggtgctc 385
Db 261 CGCTACCACCTTTGACGGAAATATGCTTCTCTGAAGACATACAACCTACTCTTTGGGTGC 320
QY 386 tgaagacttgactccattcgttgaaacaaatgggttaactctggtattaaagttctacag 445
Db 321 AGATGACCTGACTCCATTCGGAGAACAGGAGCTTGCAACTCCGGTATCAAGTTCTACCA 380
QY 446 agatacagaaggttggctagaaagattgttccattcgttagagcttctggttctgacag 505
Db 381 GAGATACGAATCTTTGACAGAAGAAATCTGTCATTTCATCAGATCCTCTGTTCTCTTAG 440
QY 506 agttattgcttctgctgaaggttcaattgaagtttccaaatctgtaagttggctgacct 565
Db 441 AGTTATCGCTCCGTTAAGAAATTCATCGAGGTTTCCAGAGCACTAAGCTCAAGGATCC 500
QY 566 aggtgctaacccacacacaggtcttccagttattaaagttatttccagaaggtgctgg 625
Db 501 TAGAGCCAGCCAGGTCAATCTTCCAAAGATCGACGTTGTCATTTCCGAGGCTCTTTC 560
QY 626 ttacaacacacttggaccaggttctgtactgtcttcgaagaattctgaagttgggtga 685
Db 561 ATCCAAACACACTCTTGACCCAGSTACTTGTACTGTCTCGAAGACTCTGAATTGGCCGA 620
QY 686 cgaagttgaagcttaacttcaactgctgttttctgctccactattagagctagattggaagc 745
Db 621 TACTGTGGAAGCCAAATTCATGCCACTTTCGTCCCATCCATTAGACAAAGACTGGAGAA 680
QY 746 tcaactgcaaggtttaacttgactgacgaagacgttgttaacttgatgacacatgtgctc 805
Db 681 CGACTGTCCGGTGTACTCTTACTGACACTGAAAGTTACTTACCTTATGACATGTGTTTC 740
QY 806 attgacactgttctagaactctgacgtactcaattgtctccattctgtgactgtt 865
Db 741 CTTCGACACTATCTCCACTTCTACCGTCGACACCAAGCTGTCCCATTTCTGTGACCTGTT 800
QY 866 cactcagcagaattggaattcaatcacgactacttgcaattcttgggtaagtaactacagtta 925
Db 801 CACCCATGAGCAATGGATCAACTACGACTACTTGCACTCTTGAAGAAAGTATTACGGTCA 860
QY 926 cgggtggttaacccattgggtccagctcaaggtgtgttggttttaaagaaattgattgc 985
Db 861 TGGTGAGGTAAACCATTTGGGTCCAAACCCAGGGTGTGCGTTACGCTTAACGAGTTGATCC 920
QY 986 tagattgactcactctccagtttaagaccacactctacttaacacactttggactctaa 1045
Db 921 CAGACTGACCCACTCTCCGTCCACCATGACACCTCTTCCAAACACACTTTGGACTCTTC 980
QY 1046 ccagctactttccactgaacgctactttgtactgctgacttctccacgaacacactat 1105
Db 981 TCAGCTACTCTTCCATTGAACCTACTTGTGACGTGACTTTTCTCATGACAAACGGTAT 1040
QY 1106 ggtttctatttctcgtcttgggttgggttgggttgggttgggttgggttgggttgggttgg 1165
Db 1041 CATCTCCATTTTGTGTTAGGTCTGTACCAACGGTACTAAGCCATTCTGTACCACACTAC 1100

```

[illegible]

3	RESULT	10-JUN-1994
	A19452	

LOCUS	DEFINITION	AL9452	1404 bp	IIIKNA
	phytase cDNA fragment.			

DEFINITION	PRI	ACC
ACCESSION	A19452	ST-503105

VERSION
KEYWORDS

KEYWORDS
SOURCE
synthetic construct.
synthetic construct.
synthetic construct.

ORGANISM
synthetic construct
artificial sequence.

REFERENCE

1 (bases 1 to 1404)

van Hartingsveldt, W., van Paridon, P.A.,

AUTHORS
van GORCOM, R. F. M., van
veenstra, A. E., Luiten, R. G. M. and Seltens, G. C. M.

TITLE Cloning and expression of microbial phylase
Patent: EP 0420358-A 41 03-APR-1991;

JOURNAL
FACILIT. ET OUVRES
GIST-BROCADES N.V.

FEATURES	Location/Qualifiers
1	1404

```

source
  /organism="synthetic construct"
  /strain="32620"

```

```

/db_xref="Taxon:32030
1.:1404
gene

```

gene
/gene="Phytase gene"
1 1404

```
CDS
L. :L404
/gene="Phytase gene"
```

```

/codon_start=1
'atrac' + chlon11

```

```

/transl_labile-11
/protein_id="CAA01464.1"

```

```

/db_xref="GI:583196"
/translation="MGVSAVLLPLYLSGVTSGLAVPASRNQSSCDTVDQGY

```

SHLWGQYAPFFSLANESVISPEVPAGCRVTFQAQLSRHGARYPTDSKGGKYY
/C/translation

QQNATTFDCKYAF LKTYNYS LGADBLTFFGEQELVNSGIRFIQRIESTLNR
SCSSBPVIASCCKKEIEGFQSTKLKDPRAQPGQSSPKIDVWISSESSSNTLDE

EDSELADTVEANFTATFVPSIRQRLNDLSGVTLTDEVTYLMDCSEDTIS

KLSPFCDLFTHDEWINYDYLQSLRKYRGAGNPLGFIQGVSIANEELIAND
DTSCSNHTIDCSSPATFPLNSTLYADFSDHNGIISILFALGLYNGTKPLSTTV

DGFSSAWTVPFASRLYVEMMQCQAEQEPLVRVLVNDVRVPLHGCVPDALGR
 DISNNHLEDCRRYFEMNPLVNDVRVPLHGCVPDALGR

	RG	LS	FAR	SGD	WAE	CFA"
202 a				436 c		344 g
202 b						331 t

BASE COUNT	293 a	430 c	430 d
ORIGIN			

Query Match	43.68;	Score 622.2;	DB 9;	Length 1404;
	5.88	2.33	50-154	

Best Local Similarity	65.2%	Pred. NO. 1.5e-154,	Indels 0;
Matches 915.	Conservative	0; Mismatches 488;	Gap 0;

Matthies 3157, Conservation 7

[illegible]

db
1 ATGGGGCTCTGCTGTCTACTTCTGTATCTCTGCTGAGTCACCTTCGGAC

72 acctgggtcctcgtggttaattctcactcttgtgacactgttgacggtgttaccaatgt I

12345678910111213141516171819202122232425262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989910010110210310410510610710810911011111211311411511611711811912012112212312412512612712812913013113213313413513613713813914014114214314414514614714814915015115215315415515615715815916016116216316416516616716816917017117217317417517617717817918018118218318418518618718818919019119219319419519619719819920020120220320420520620720820921021121221321421521621721821922022122222322422522622722822923023123223323423523623723823924024124224324424524624724824925025125225325425525625725825926026126226326426526626726826927027127227327427527627727827928028128228328428528628728828929029129229329429529629729829930030130230330430530630730830931031131231331431531631731831932032132232332432532632732832933033133233333433533633733833934034134234334434534634734834935035135235335435535635735835936036136236336436536636736836937037137237337437537637737837938038138238338438538638738838939039139239339439539639739839940040140240340440540640740840941041141241341441541641741841942042142242342442542642742842943043143243343443543643743843944044144244344444544644744844945045145245345445545645745845946046146246346446546646746846947047147247347447547647747847948048148248348448548648748848949049149249349449549649749849950050150250350450550650750850951051151251351451551651751851952052152252352452552652752852953053153253353453553653753853954054154254354454554654754854955055155255355455555655755855956056156256356456556656756856957057157257357457557657757857958058158258358458558658758858959059159259359459559659759859960060160260360460560660760860961061161261361461561661761861962062162262362462562662762862963063163263363463563663763863964064164264364464564664764864965065165265365465565665765865966066166266366466566666766866967067167267367467567667767867968068168268368468568668768868969069169269369469569669769869970070170270370470570670770870971071171271371471571671771871972072172272372472572672772872973073173273373473573673773873974074174274374474574674774874975075175275375475575675775875976076176276376476576676776876977077177277377477577677777877978078178278378478578678778878979079179279379479579679779879980080180280380480580680780880981081181281381481581681781881982082182282382482582682782882983083183283383483583683783883984084184284384484584684784884985085185285385485585685785885986086186286386486586686786886987087187287387487587687787887988088188288388488588688788888989089189289389489589689789889990090190290390490590690790890991091191291391491591691791891992092192292392492592692792892993093193293393493593693793893994094194294394494594694794894995095195295395495595695795895996096196296396496596696796896997097197297397497597697797897998098198298398498598698798898999099199299399499599699799899910001001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110111021103110411051106110711081109111011111112111311141115111611171118111911201121112211231124112511261127112811291130113111321133113411351136113711381139114011411142114311441145114611471148114911501151115211531154115511561157115811591160116111621163116411651166116711681169117011711172117311741175117611771178117911801181118211831184118511861187118811891190119111921193119411951196119711981199120012011202120312041205120612071208120912101211121212131214121512161217121812191220122112221223122412251226122712281229123012311232123312341235123612371238123912401241124212431244124512461247124812491250125112521253125412551256125712581259126012611262126312641265126612671268126912701271127212731274127512761277127812791280128112821283128412851286128712881289129012911292129312941295129612971298129913

Db
61 GCAGTCCCGCTCGAGAAATCATTCGATCCGTTGGTTT

132	ttccagaaattctcaacttgggggtcaataactctccattctctcttgggtgaagaa	191
Qy		
Db	TTCTCCGAGACTTCGCATTTTGGGTCAATACGACCGCTCTCTGTGTGGCAACGAA	180
Qy		
Db	tctgtctatttctcagacgttccaaaggggtgtagaattactcttgcgtcaagtgttgc	251
Qy		
Db	181 TCGGTCAATCTCCCGCTGAGGTGCCCGCGATGCGAGATCACTTTTCGTCAGGTCTCTCC	240
Qy		
Db	252 agacaggtctagataccaaactcttcttaagttctaaagtlactctgtctttagaa	311
Qy		
Db	241 GGTATGAGGGCGGTATCGGACGACTCCAAGGGCAAGAAATACCTCCGCTCAATTGAG	300
Qy		
Db	312 gctattcaaaagaacgctactgcttccaagggtaagtacgctttcttgaagacttacaac	371
Qy		
Db	301 GAGATCCACGACGACGACACCTTTGACGGAATATGCTTCTCGAGACATACAC	360
Qy		
Db	372 tacacttgggtgctgacgacttgcactcattcggtgacaacaaaatgggttaactctgt	431
Qy		
Db	361 TACACCTTGGGTGCAGATGACCTGACTCCGCTTCGGAGAACAGGAGTAGTCAACTCCGCG	420
Qy		
Db	432 attaggttctcagaagatacaaggttggctgataaagattgtccattcgttagagct	491
Qy		
Db	421 ATCAAAGTTCTACGACGGTACGAAATGCTCACAGGACATCGTTCATTATCCGATCC	480
Qy		
Db	492 tctgggtctgacagagtattgtctgtgtgaaaagttcaattgaaggtttccaatctgct	551
Qy		
Db	481 TCTGGCTCAGCGCGGTGATCGGCTCCGGCAGAAATTCATCGAGGGCTTCCAGCACCC	540
Qy		
Db	552 aagttggctgacacaggtctaaacccacacaaggtctccagttattaaagcttattt	611
Qy		
Db	541 AAGCTGAAGGATCTCTGTCGCCACCGCGGCATTCGTGCCCAACATCGAGTGGTCAAT	600
Qy		
Db	612 ccagaaggtgtggtttacaacaaactttggacacaggttgtgtactgctttcgaagaa	671
Qy		
Db	601 TCCGAGGCCAGCTCATCTCAACAAACACTCTCGACCGAGGACCTGACACTGCTTCGANGAC	660
Qy		
Db	672 tctgaattgggtgacgagcttgaagctaaactcactgctgttttgcgtccaccctattaga	731
Qy		
Db	661 AGCGAATGGCCGATACGCTCGAAGCCAAATTCACGCCACGTTGTGCTCCCTCCATCGT	720
Qy		
Db	732 gctagattggaagctcaacttgcagaggtttaactgaactgacgaagaagttgtaacttg	791
Qy		
Db	721 CAACGCTCTGAGAACGACCTGTCCGGTGTGACTCTACAGACACAGAGTGACCTACCTC	780
Qy		
Db	792 atgacatgtgtccattcgacactgttgcgtagaactctgacgctactcaattctcca	851
Qy		
Db	781 ATGGACATGTCTCTCGACACCACTCTCCACGACCGCTGACACCAAGCTGTCCCCC	840
Qy		
Db	852 ttctgtaactgttcaactcagcaagaatggattcaatacgaactacttgaactctttgggt	911
Qy		
Db	841 TTCTGTGACCTGTGTCACCATGACGANTGGATCAACTACGACTACCTCCAGTCTTGAAA	900
Qy		
Db	912 aagtlactacgtttacggtgtgtgaaccocatlygggtccagctcaaggtgttggttcggt	971
Qy		
Db	901 AAGTATTACGGCCATGTGTGAGGTACCGCTTCGCGCCGACCCAGGCGTTCGGCTACGCT	960
Qy		
Db	972 aacgaattgattcgtagattgactcaactctccagttcaagaccactcttactaacac	1031
Qy		
Db	961 AACGAGCTCATCGCCCGCTCTGACCCACTCGCTGCTCCAGATGACACGATTCCCAACAC	1020
Qy		
Db	1032 actttggactctaaccagactaatttcccatgaaagcgtactttgtacgctgactctct	1091
Qy		
Db	1021 ACTTTGGACTCGAGCGCGGTACTCTTCGCGTCAACTCTACTCTCTACGGGGACTTTTCG	1080
Qy		
Db	1092 cagcaacaacataagtttcttatttcttgcgttttgggtttgtgacaacgtaactaagcca	1151
Qy		
Db	1081 CATGACACGCGCATCTCCATTCCTTTGGTTTAAAGTCTGTGTACAGCGCACTAAGCG	1140
Qy		
Db	1152 ttgtctactactctctgtgaaactattgaagaaactgacggttgcagctgcttcttgact	1211
Qy		
Db	1141 CTATCTTACCACAGACCGTGGAGAAATCACCCACAGACATGGAATCTCTGCTCTGTGGACG	1200
Qy		

100

Sat Oct 27 15:25:36 2001

SPFCOLFTHEWKYNYLQSLGKYGAGNPLPAGQIGFTNELIARLSPQDHT
 STNSLVSNPATFLNATMYDFSHDNMSWIFFALGLYDFNLEPLSRTSVESAKELDG
 YSASWVPFGARAYETWOCKSEKPELVRALINDRVPLHGCDDVDKLGKRLKDNFVKG
 LSWRSGGNWGCPS"
 BASE COUNT 377 a 480 c 532 g 423 t
 ORIGIN

Query Match 43.5%; Score 619.6; DB 15; Length 1812;
 Best Local Similarity 67.0%; Pred. No. 7.5e-154;
 Matches 896; Conservative 0; Mismatches 439; Indels 3; Gaps 1;

Qy 86 tggtaattctcaactcttggtagactgttgacaggttgagcgggtttaccacatgtttcccaagaatttc 145
 Db 412 TCGTGGCTCCCAAGTCCCGCATACCGTACGCTAGACCTCGGGTACCGAGTGTCCCGTGCAG 471
 Qy 146 tcaacttgagggtcaataactctcattctctctcttcttggctgacgaattctgctattctcc 205
 Db 472 TCATCTATGGGGCCAGTACTCGCCATCTTTTCGTCGAGGAGAGTGTCCCGTGCAG 531
 Qy 206 agacgttccaaagggttgtagacttcttcttcttcttcttcttcttcttcttcttcttcttct 265
 Db 532 TAACTCTCCCAAGGATTCGCGATACCTTGGTACAGGTGCTAFCGCGCATGAGCGCG 591
 Qy 266 ataccacacttcttaagcttaagaagctactctgttcttcttcttcttcttcttcttcttctt 325
 Db 592 GTACCCACACGCTCCAGAGCAAAAAGTATAAGAACCTTGTGACGCGCATCCAGGCCAA 651
 Qy 326 cgtactgttttcaagggttaagctacttcttcttcttcttcttcttcttcttcttcttcttct 385
 Db 652 TGCACCGACTTCAAGGCAAGCTTTGGCTTTTGAAGACGTTACAACTATATCTCFGGGTGC 711
 Qy 386 tgacgacttgactcaattcttcttcttcttcttcttcttcttcttcttcttcttcttcttctcag 445
 Db 712 GGATGACCTCACTCCCTTTGGGGAGCAGCAGCTGTGAACCTCGGCATCAAGTCTACCA 771
 Qy 446 aagatacaagggttggcttagaagattgttccattcttcttcttcttcttcttcttcttcttct 505
 Db 772 GAGGTACAGGCTCTGCGCGAGTGTGTGGCTTTATTCGCGCTCAGGCTCGGAGCG 831
 Qy 506 agttattgtctgtgaaagtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctc 565
 Db 832 GGTATTGCTTCGGGAGAGAGTTCATCAGGGGTTCACAGCGGAGAGCTGGCTGATCC 891
 Qy 566 agdtgttaacccacacacagcttcttcttcttcttcttcttcttcttcttcttcttcttctt 625
 Db 892 TGGCGCA---CGAACCGCGCTCCGCGATTAGTGTGATTATTCGCGAGAGCGAGAC 948
 Qy 626 ttacaacacacacttggacacacggttcttcttcttcttcttcttcttcttcttcttcttctt 685
 Db 948 GTTCAACAATACGCTGACACCGGTGTGTGCACCAAGTTCAGGCGAGTTCAGCTGGAGA 1008
 Qy 686 cgagctgaagcttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctc 745
 Db 1008 TGAGTTTCGGGCCAATTCCTACTGGCTCTTTCACCGCATATCCGAGCTCGCGCGAGAA 1068
 Qy 746 tcaacttgccaggttgaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 805
 Db 1068 GCATCTTCTGGCGTACGCTGACAGACGAGGACCTTCTCAGTCTAATGGACATGTGTTTC 1128
 Qy 806 attcagactgttctgaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 865
 Db 1128 GTTTGATAGGTAGGCGGCACACGACGCAAGTACAGTCTGTACCGCTTCTGTCAACTCTTT 1188
 Qy 866 cactcagcagcaatggttcaatacagactcttcttcttcttcttcttcttcttcttcttcttctt 925
 Db 1188 CACTCACAATGAGTGAAGAGATCAACACTCTTCAGTCTTCTGGCAAGTACTACGSGCTA 1248
 Qy 926 cgttctgttaacccattgttccagcttcaagggttcttcttcttcttcttcttcttcttcttct 985
 Db 1249 CGCGGACGACACCTCTTGGGACCGGCTCAGGGGATAGGTTTCAACCAACGAGCTGATTGC 1308
 Qy 986 tagattgactcactctccaggttcaagacacacacttcttactaaccacttcttggactctaa 1045

RESULT 8

ANPHYAG ANPHYAG 2000 bp DNA PLN 06-OCT-1992
 LOCUS A.niger phyA gene.
 DEFINITION Z16414
 ACCESSION Z16414.1 GI:2392
 VERSION phyA gene.
 KEYWORDS Aspergillus niger.
 SOURCE Aspergillus niger.
 ORGANISM Aspergillus niger.
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE 1 (bases 1 to 2000)
 AUTHORS Van Hartingsveldt, W., Van Zeijl, C. M. J., Harteveld, M. G., Gouka, R. J.,
 Suykerbuyk, M. E. G., Luiten, R. G. M., Van Paridon, P. A., Sellen, G. C. M.,
 Veenstra, A. E., Van Gorcom, R. F. M. and Van Den Hondel, C. A. M. J.
 Cloning, molecular characterization and overexpression of the
 phytase gene (phyA) of Aspergillus niger
 JOURNAL Gene (1992) In press
 REFERENCE 2 (bases 1 to 2000)
 AUTHORS van Hartingsveldt, W.
 TITLE Direct Submission
 JOURNAL Submitted (05-OCT-1992) Van Hartingsveldt W., TNO Medical
 Biological Laboratory, Lange Kleiweg 139, Rijswijk, the Netherlands
 FEATURES
 source
 1. .2000
 /organism="Aspergillus niger"
 /strain="Van Tiegheem strain"
 /db_xref="taxon:5061"
 TATA_signal 131. .134
 exon 166. .253
 /gene="phyA"
 /number=1
 166. .1871
 /gene="phyA"
 prim_transcript 166. .1871
 /partial
 /gene="phyA"
 /function="primary transcript of phyA gene"
 join(210. .253, 356. .365)
 /gene="phyA"
 join(210. .253, 356. .1715)
 CDS

/gene="phyA"
 /codon_start=1
 /protein_id="CAA78904.1"
 /db_xref="GI:2393"
 /translation="MGVSALLPLLYLLSGVTSGLAVPASRNSSDDVDVQGYOCFSET
 SHWGOYAPFPFLSANEVSISPEVPACRVTPAQVLSRHGARYPTDSKGYKYSALIEEI
 QONATTFDQKAYFLKTVNYSLGADDTLPFGQELVNSGKIFORYESLFRNIVPTIRS
 SSSRVIASCKFIEGFSQTKLKDPAQPGSQSPKIDVYISASSNPTLDPTCTVF
 EDSLEADTVENAFITVPFISIRORLENDLSGVTLDTTEVYLMDCMSPTDSTSTVDI
 KLSFCDLTHDEWINYDLSKLYYGHGANGPLGPTQGVYANELARLTHSPVHD
 DYSNHTLSDSPATFPLNLSVADFSHDNGIISILFALGLYNGTLPKLSITVENITOT
 DGFSSANTVPFASRLYVENMQCAQEPFLVRLVNDVRVPLHGCPVDALGROTRDSFV
 RGLSFARSGCDWAECFA"
 254..355
 intron
 /gene="phyA"
 /number=1
 356..1878
 exon
 /number=2
 369..1712
 mat_peptide
 /gene="phyA"
 1732..1760
 stem_loop
 /gene="phyA"
 463 a 603 c 456 g 478 t
 BASE COUNT
 ORIGIN

Query Match 43.4%; Score 618.2; DB 13; Length 2000;
 Best Local Similarity 65.3%; Pred. No. 1.8e-153;
 Matches 908; Conservative 0; Mismatches 483; Indels 0; Gaps 0;

QY 24 gtcgtgactgtccatgcccaccttggctcggtccacatccggtacgcgcttgggtctc 83
 Db 324 GTGGGACTACTGATCGCTGACTATCTGTGAGAGTCACCTCCGAGTGGCAGTCCCGGCC 383
 QY 84 cgtgtgaattcactctgtgacactgtgtgacggtgtgtaccactgtttcccaagaatt 143
 Db 384 TCGAGAAATCAATACAGTTCGATACGCTGATCAGGGGTATCAATGCTTCTCCGAGACT 443
 QY 144 tctcaacttgggttcaactactcctccattctcttcttcttggctgacgaactgtctattct 203
 Db 444 TCGCATCTTTGGGGTCAATACCGCACCGCTTCTTCTCTGCTCAACAGGAATCGGTCACTCC 503
 QY 204 ccagacgttccaaaggtgttagagttacttctcggttcaagtttggctgacgacggtgct 263
 Db 504 CCTGAGTGGCCGCGGATCCAGAGTCACATTCGCTCAGGTCCTCTCCCGCTCATGGAGCG 563
 QY 264 agatacccaacttcttaagttaagaagtactctgtctgttgaattgaagctattcgaaga 323
 Db 564 CGGTATCCGACCGACTCCAAAGGCAAGAAATACTCCGCTCTCATTGAGGAGATCCAGCAG 623
 QY 324 aacgtactgtcttcaaggttaagtaagcttcttcttgaagacttacaactacacttgggt 383
 Db 624 AACCGGACCACTTTGACGGAAATATGCGTCTCTGCTCAAGACATACAACTACAGCTGGGT 683
 QY 384 gctgacgactgactccattcgtgtaacaaatgttgaactctggtattgaagtcttac 443
 Db 684 GCAGATGACGTGACTCCCTTCGAGAACAGGAGCTAGTCAACTCCGCGCATCAAGTCTTAC 743
 QY 444 agaagatacaaggttggctgagaagattgttccattcgttagagcttctggttctgac 503
 Db 744 CACCGGTAGAAATCGCTCAAGAGGAACATCGTTCCATTATCCGATCTCTGGCTCCAGC 803
 QY 504 agagttattgcttctgtaaggttcaattgaaggtttccaatctgctaagttgggtgac 563
 Db 804 CGCGTGATCGCTCCGGAGAAATATCATCGAGGCTTCCAGAGCACCAGCAAGGAT 863
 QY 564 ccaggtgtaaccacacaaagcttctccagttatttaacgcttatttccagaaggtgct 623
 Db 864 CTCCTGCGCCAGCCGCGCAATCGTCGCCCAAGATCGAGCTGGTGTATTTCCGAGGCCAGC 923
 QY 624 ggttacaacaacacttggaccaggttgggtgactgcttctcgaagaatctgaattgggt 683

Db 924 TCATCCAAACAACACTCTCGACCGACGCACTGCACCTGTCTTGAAGACAGCAAGTTGGCC 983
 QY 684 gacgacgttgaagcttaacttcaactctgtgttttcgctccactattagagctagattgaa 743
 Db 984 GATACCGTGAAGCAATTTACCCCGAGGTTCCGTCCTCCCTCCATTTCGTAACGCTCTGGAG 1043
 QY 744 gctcaactgcccaggtgttaacttgaactgacgaagcgttgttaacttgaatgacatgct 803
 Db 1044 AACGACCTGTCCGGTGTGACTCTCAGACACAGAGTGACCTACCTCATGACATGACATGTC 1103
 QY 804 ccattcgcactgttctgtagaacttctgacactactgacactcaattgttccattctgtgactg 863
 Db 1104 TCCTTCGACACCATCTCCACGACGACCGCTCGACACCAAGCTCTCCCTCTCTGTGACCTG 1163
 QY 864 ttaactcagcgaatggattcaatcagcactacttgcacttgcacttcttgggttaagctacggt 923
 Db 1164 TTCACCATGACGAATGATCACTACGACTACCTTCCAGTCTCTGAAAAAGATTATACGGC 1223
 QY 924 taaggtgctgtaaccacttgggtccagctcaaggtgtgtgttctgttcaacgaattgatt 983
 Db 1224 CATGTCGAGGTAACCGCTCGGCCGACCCAGCCAGGCGCTCGGCTACGCTAACGAGCTCATC 1283
 QY 984 gctagattgactcactctccagttccaagacacacttctactaacacacttggactct 1043
 Db 1284 GCCGCTGACGACCTCGCCTGTCCACGATGACACACAGTTCCAAACACACTTTTGGACTCG 1343
 QY 1044 aaccagctacttccattgaacgctacttgtacgctgacttctcactcaacacact 1103
 Db 1344 AGCCGCTACTCTTCGGCTCACTCTACTCTCTACCGGACTTTTCGATGATCAACAGGC 1403
 QY 1104 atggttcttatttctcgttgggtttgttgaacaggttacttaagcattgtctactact 1163
 Db 1404 ATCATCTCCATTCTTGTCTTGTAGTCTGTACACGGCCTAACGCCGTATCTACACAG 1463
 QY 1164 tctgtgaattcttattgaagaactcagcgttcaagctgttcaagctgttccacttccact 1223
 Db 1464 ACGGTGAGNAATATACCCAGACAGATGATGATTCCTCTCTGCTGGACGCTCCGTTGCT 1523
 QY 1224 gctagacttaagttgaatgatgcaatgtgaagctgaaagaaacattggttagagtt 1283
 Db 1524 TCGCGTTTCTAGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
 QY 1284 ttggttaacacagaggttctccattgcaaggttgggttgggttgggttagagtt 1343
 Db 1584 TTGGTTAATGATCGCGGTTGTCCCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1643
 QY 1344 aagagacgacttctggtgaaggttgggttgggttgggttgggttgggttgggttgggttgggt 1403
 Db 1644 ACCCGGATAGCTTTGTCAGGGGTTGAGCTTTGCTAGATCTGGGGTGTATTGGCGGAG 1703
 QY 1404 ttttctgctta 1414
 Db 1704 TGTTTTGCTTA 1714

RESULT 9
 AR018076
 LOCUS
 DEFINITION Sequence 7 from patent US 5780292.
 ACCESSION AR018076
 VERSION AR018076.1 GI:3973679
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2363)
 AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
 Torckeli,T.K., Cantrell,M., Piddington,C.S., Rambosek,J.A.,
 Turunen,M.K. and Fagerstrom,R.B.
 TITLE Production of phytate degrading enzymes in trichoderma
 JOURNAL Patent: US 5780292-A 7 14-JUL-1998;
 FEATURES Location/Qualifiers
 source 1..2363

[illegible][illegible]

BASE COUNT	559 a	732 c	510 g	562 t	
ORIGIN	/organism="unknown"				
Query Match	43.48;	Score	618.2;	DB 9;	Length 2363;
Best Local Similarity	65.3%;	Pred. No.	1.8e-153;		
Matches 908;	Conservative	0;	Mismatches 483;	Indels	0;
Gaps					
QY	24	gtcgtgtaactgcatctgacacattgttgcggttccacatcccggtacgcgcttgggtctc	83		
DB	518	GTGGGACTACTGATCGCTGACAAATCTGTGCAGAGTCACTCCGGACTGGCAGTCCCGCCG	577		
QY	84	ctgtgtaattcactcttctgtgacactgtgacggtgttaccagaatttccacagaatt	143		
DB	578	TCGAGAAATCAATCCACTTGGCATACGGTTCGATCAAGGGTATCAATGCTTCTCCGAGACT	637		
QY	144	tctcacttgggggtcaatcactctccattctctcttcttggctgcgacgaattctgtattct	203		
DB	638	TCCGATCTTTGGGTCAATACGCGCGTCTCTCTCTGCGCAACGAATCGGCACTCC	697		
QY	204	ccagaagttccaaagggtgtgtagtactcttgcctcaagtttctgtctagacacggtgct	263		
DB	698	CCTGATGTGCCCGCGGTGTGCAGAGTCACTTTCGCTCAGTCTCTCCCGTCATGGAGCG	757		
QY	264	agatacccaacttcttctaagttctaagaagtactctgttgaatgaactattccaaag	323		
DB	758	CGGTATCCGACCGAGTCCAAGGCGCAGAAATATCTCCGTCTCTGTAGGAGATCCAGAG	817		
QY	324	aacgtactgcttccaaagggttaagtacgcttcttgaagacttacaactacactttgggt	383		
DB	818	AACGTGACCACTTTGATGAAATATGCTTCTCTGAGACATACACTACAGCTTGGGT	877		
QY	384	qctaacgacttgactcattctcgttgacacacaaatggttaactcgtgtattaaagttctac	443		
DB	878	GGAGATGACTGTACTCCCTTCGGAGACAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC	937		
QY	444	agaagatacagaaggcttggctagaagaattgtccattggttagagcttctgttctgac	503		
DB	938	CAGCGATACGAATCGCTCACAAAGGAATCATCTCCGTTTCATCCGATCTCTGGCTCCAGC	997		
QY	504	agagttattgctctgtgaaagtctcattgaagggtttccaatctgtctaagttggtcag	563		
DB	998	CGCGTATCGCCTCGCGGAGAAATTCATTGAGGCTTCAGAGACCAAGCTGAAGAT	1057		
QY	564	ccaggtgttaaccacacacagcttcccggtattaaagttatttccagaagagtgct	623		
DB	1058	CCTGFTGCCACCGCGGCGCAATCGTCGCCCNAGATCGAGCTGGTTCATTTCCGAGGCCAG	1117		
QY	624	ggttaacaaacactttggaccagggtttggtagctgttccgaagaatctgaattgggt	683		
DB	1118	TCATCCAAACACACTCTCGACCGAGCACTGCACTCTCTTTGAAGACACGAATTTGGCC	1177		
QY	684	gacgagttgaagctaacttcaactgctgtttctgctccactattagagctagattggaa	743		
DB	1178	GATACCGTGAAGGCCAATTTACCGCCAGCTTGCCCCCTCCATTGCTCAACGCTGGAG	1237		
QY	744	gctcacttgcaggtgttaacttgaactgacgaagacggttgttaacttgatggacatgtgt	803		
DB	1238	AACGACCTGTCTGGCGTGACTCTCACAGACACAGAAGTGACTACCTCATGGACATGTGC	1297		
QY	804	ccattcgacactgttgcctagaactttgaactactcgaattgtccattctgtgacttg	863		
DB	1298	TCCCTTCGACACCACTTCTCCACAGACCGCTCGACACCAAGCTGTCCCCCTTCTGTGACCTG	1357		
QY	864	ttaactcacgacgaattgatttcaatcacgactacttgcgaatttgggttaagctactacggt	923		
DB	1358	TTACCCATGACGAATGATCCACTACGACTACCTCCAGTCCAGTCCCTGAAAAAATACATCGGC	1417		
QY	924	tacggtgtcgttaaccacttgggttccagctcgaaggttgggttcttgcgtttaaocgaattgatt	983		
DB	1418	CATGCGCAGGTAAACCGCTCGGCCGACCCAGGGCGTCGGCTACGCTAACGAGCTCATC	1477		

Db	698	CTGTATGCGCCGCGGTTTCAGAGTCACTTTCGCTCAGTCTCTCCCGTCATGGAGCG	757
QY	264	agataccaaacttcttctaagttctaagaagtactctgttgattgaagctattcaaaag	323
Db	758	CGGTATCCGACCGAGTCCRAGGCCAAGAAATCTCCGCTCTCTTTTTCATTCAGGAGATCCAGCAG	817
QY	324	aacgctactgctttcctaagggttaagtcgcttctttgaagacttaacaactacactttgggt	383
Db	818	AACGTGACCACCTTTGATGGAAATATGCTTTCCCTGAAGACATACAACTACACCTTGGGT	877
QY	384	gctgacgacttgactctcatctcggtgaacaacaaatggttaacctggttattaagttctac	443
Db	878	GCAGATGACCTGACTCCCTTCGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTCTTAC	937
QY	444	agaaatatacaaggcttggctagaaagattgttcattctggttagagctctggtttctgac	503
Db	938	CAGCGATAGCAATCGCTCACAGGAACATCATTCGGTTTCATCCGATCCTCTGGGTCCAGC	997
QY	504	agagttattgctctgctgtaaaagtctaatgaaggtttccaaatbctgaagttggctgac	563
Db	998	CGCGTGATCGCTCGGCGCGAATTCATTGAGGCTCTCCAGAGCACCAAGTCAAGGAT	1057
QY	564	caaggtgtaacccacacacccaagctctccagttattaaagttatttcacagaagtgct	623
Db	1058	CCTGTGCCAGCGCGGCGCAATCTGCGCCCAAGATCGAGCTGGTCAATTCGAGGCCGACG	1117
QY	624	ggttacacaacacttggaccacaggttgggtactcttcgaagaatctgaattgggt	683
Db	1118	TCATCCAAACACACTCTCGACCCAGGCACCTGCACTCTTTGAAGACACGCAATTGGCC	1177
QY	684	gacgactgtaagcttaacttcaactgctgttttcgctccacctattagagctagatggaa	743
Db	1178	GATACCTCGAAGCCAAATTCACGGCACCTTCGCCCCCTCCATTCTCAACGCTCTGAG	1237
QY	744	gctcaactgccaagtttaacttgactgaagaagacgttgttaacttgatggacatggt	803
Db	1238	AACGACCTGTCTGGGTGACTCTACAGACACAGAAGTGACCTACCTATGGACATGTGC	1297
QY	804	ccattgacaactgtgtgtagaactctcgactcaactcaattgctcattctgacttg	863
Db	1298	TCCTTGTGACACCATCTCCACGACCGCTGCACACCAAGCTGTCCGCTCTGTGACCTG	1357
QY	864	ttcaactcaagcaaatggaattcaatacgaactacttgcaatcttggglaagtaactacgt	923
Db	1358	TTACCCCATGACAAATGATGCATACGACTACCTCCAGTCCCTGAAATAATACTACGGC	1417
QY	924	taogtgtggttaaaccaattgggtccagctcaaggtgtgttggttaacgaattgatt	983
Db	1418	CATGGCGCAGGTAAACCGCTCGCGCCGACCCAGAGGCGTCCGGTACGCTAACGAGCTATC	1477
QY	984	gctagttgactcaactctcagttcaagacacacttctactaaccacactttggactct	1043
Db	1478	GCCGCTCTACCCACTCGCGTGTCCACGATGACACAGCTCCACACACACTTGGACTCG	1537
QY	1044	aaccagactcttcccattgaagctactttgacgtgactttgacgtgactctctcaagcaacact	1103
Db	1538	AACCCAGCTACCTTCCCGCTCAACTCTACTCTCTACGGCGGACTTTTCCCAGATTAACGGC	1597
QY	1104	atggttctatttctctogcttgggttttgacaacggtactaagccattgcttactact	1163
Db	1598	ATCATCTCTATTCCTTTTGCTTTGGGCTGTGTACAACGGCACTAAGCCGCTGTCTACCAAG	1657
QY	1164	tctgtgaactattgaagaaactgacggttagcgtgcttcttggaactgttccattcgt	1223
Db	1658	ACCGTGGAGAATATCACCCAGACAGATGGGTTCTGCTCTGCTTGGAGGTTTCGGTTTGT	1717
QY	1224	gctagagcttacgttgaaatgatcgaatgtgaagctgaaaggaacccattggtagatt	1283
Db	1718	TCGGCTGTGATCGAGATCATCGAGTGCAGCGCGAGCAGGAGCGCGTGTGCCGTGTC	1777
QY	1284	ttggttaacgacagagttgttccattgcacggttgtggtgttgacaagttggtagatgt	1343

Db	1778	TTGGTAAATGATCGCTGTGTCCGCTGCATGGGTCCCAATTGATGCTTTGGGGAGATGT	1837
Qy	1344	aagagagacgacttcgttgaaggttcttcctcgtatgatcttggtgtaacttgggaagaa	1403
Db	1838	ACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGTGTATGGCGGAG	1897
Qy	1404	tgtttcgctta 1414	
Db	1898	tgtttcgctta 1908	
RESULT	11		
LOCUS	AR053934		
DEFINITION	Sequence 1 from patent US 5834286.	DNA	29-SEP-1999
ACCESSION	AR053934		
VERSION	AR053934.1	GI:5978796	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2379)		
AUTHORS	Mevalainen,H.K.M., Paloheimo,M.T., Fagerstrom,R.B., Niittinen-Oinonen,A.S.K., Turunen,M.K., Rambosek,J.A., Piddington,C.S., Houston,C.S. and Cantrell,M.A.		
TITLE	Recombinant cells that express phytate degrading enzymes in desired ratios		
JOURNAL	Patent: US 5834286-A 1 10-NOV-1998;		
FEATURES	Location/Qualifiers		
Source	1..2379		
BASE COUNT	561 a 735 c 518 g	565 t	
ORIGIN			
Query Match	43.4%	Score 618.2; DB 9; Length 2379;	
Best Local Similarity	65.3%	Pred. No. 1.8e-153;	
Matches	908; Conservative	0; Mismatches 483; Indels	0; Gaps 0;
Qy	24	gtcgtgctactgtccattgccacctgttcggttcacatccggtaccgcttgggtcct	83
Db	534	GTGGGACHTACTGCTGACAACTGTGTCAGAGTCACCTCCGGACTTGGCAGTCCCGCC	593
Qy	84	cgtggaattctcactgttgacactgttgacggtgttaccaatgtttcccaagaatt	143
Db	594	TGGAGAAATCAATCCACTTGGCATACGTGTCATCAAGGGTATCAATGCTTCTCCGAGACT	653
Qy	144	tctcaactgtgggtcctaactctccattctcttcttggctgacgaattctgtattct	203
Db	654	TGCGATCTTTGGGTCATAACGCGCGGCTTCTCTCTGGCAACGAATCGGCCATCTCC	713
Qy	204	ccagagcttccaaaggtgttagattactttcgttcaagtttgtctagacacagtgt	263
Db	714	CCTGATGTGCGCGCGGTTGCAGAGTCACCTTCGCTCAGTCTCTCCCGTCATGAGGG	773
Qy	264	agatacccaacttcttaagtctaagaagtactctgtttgattgaagctattcaaaag	323
Db	774	CGGTATCCGACCGAGTCCAGGCGCAAGAATACTCCGCTCTCAATTGAGGAGATCCAGCAG	833
Qy	324	aagctactgctttoaaggtgaagtacgcttcttgaagaacttcaacttactttgggt	383
Db	834	AAGCTGACCACTTTGATGGAAATATGCTTCTTGAAGACATACACTACAGCTTGGGT	893
Qy	384	gctgaagcttgaactccattcgtgaaacaaaatgggttaactcttggattcaagttctac	443
Db	894	GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC	953
Qy	444	agaagatacaaggtcttggctgaagaagtgttccattcgttagagcttctgggtctgac	503
Db	954	CAGCGATCGAATCGCTCAACAGGAACATCATTCGTTTATCCGATCTCTGGCTCCAGC	1013
Qy	504	agagttatgcttcgtgaaagttcattgaagtttccaatctgtaagttggtgac	563

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	FEATURES	source
1014	CGCGTGATCGCCTCGCGGGAATAATTCATTGAGGGCTTCCAGAGCACCAAGCTGAAGGAT	1073				
564	ccaggtgtaaccacacacacagcttctccagttattataagctattattccagaagtgct	623				
1074	CTCGTGCCAGCGGGCCAAATCTGCCCAAGATCGAGGTGTCATTTCCGAGGCCAGC	1133				
624	ggttacacacacacacacacaggttltgtactgcttctcgaagaatctgaattggat	683				
1134	TCATCCAAACACACTCTCGACCCAGGACCTGCAGCTGCTTTGAAGACAGCGAAATGGCC	1193				
684	gacgagcttgaaagcttaactctgctgttttctgctccacactattagagctagatggaa	743				
1194	GATACCGTCGAAGCAATTCACGCGCCAGCTTCGCCGCCCTCCATTCGTCAACGTCGAG	1253				
744	gctcactgacaggtgtaactctgactgaagaacagctgtgttaactgtatggacatgtgt	803				
1254	AACGACCTCTCTGGCGTGACTCTCACACACACAGAGTGACCTACCTCATGGACATGTGC	1313				
804	ccattgcacactgttctagaaactctgacgctactcaaatgtgtctcactctgtgacttg	863				
1314	TCCTTCGACACCATCTCCACCAGCAGCTCGACACCAAGCTGCCCTCTCTGACCTG	1373				
864	ttcaactcaagagaatgaattcaatacactactctgcaactcttgggttaagctacaggt	923				
1374	TTCAACCATGACGAATGGATTCACCTACGACTACCTACCTCCAGTCCCTGAAAAAATAC	1433				
924	tacggtgtgtgaaccattgggtccagctcaagctgaagtggtgtgtttgttaacgaattgatt	983				
1434	CATGGCGCAGGTAAACCGCTTCGCGCCGACCCAGGCGTGGCTAGCTTAACGAGCTCATC	1493				
984	gctagatgactcaactctccagttcaagacacacactcttactaaacacacactttggaactct	1043				
1494	GCCGCTCTCACACACTCCCTCTGACGATGACACACAGCTCCAAACACACCTTGGACTCG	1553				
1044	aaccagactcttcccaactgaagcgtactctgtacgctgactctctcgaagacacact	1103				
1554	AACCCAGCTACTCTCCCGCTCAACTCTACTCTCTACGCGGACTTTTCCACACATTAACGC	1613				
1104	atggttctattctctgcttcttgggtttgtaaacagcgtactaagcattgtcttactact	1163				
1614	ATCAFCCTACTCTCTTGGCTTGGCTCTGTACACGGCACTAAGCGCGTGTCTACCAAG	1673				
1164	tctgttgatctatgaagaacacagcgttaagctgcttcttggactgttccattcgt	1223				
1674	ACCGTGGAGATATACCCACACAGATGGGTTCCTGCTCTGCTTGGACGGTTCGGTTGCT	1733				
1283	gctagagctacgttgaaatgatcaatgtgaagctgaaaggaacacattggttagatt	1283				
1734	TCGCGCTCTGACGATGATGACGATGCGCAGCGCAGGAGGCGCTGGTCCGTGTC	1793				
1284	ttggttaacacagaggtttccattgcacaggttgggtgtgttgacaaagttggtagatt	1343				
1794	TGGTTAATGATCGGTTGTCGCCGTCGATGGGTGTCGAATGTATGCTTTGGGAGATGT	1853				
1344	asagagacacactggttgaaggttcttctcgtctagatctggtgtgaactgggaagaa	1403				
1854	ACCCGGGATGAGCTTTGTGAGGGGTTGACCTTCTGCTAGATCTGGGGGTGATGGCGGAG	1913				
1404	tgtttcgctta	1414				
1914	TGTTCTGCTTA	1924				
RESULT	12					
ASNPHYTAS						
LOCUS	ASNPHYTAS	2379 bp	DNA	PLN	23-MAR-1994	
DEFINITION	Aspergillus niger var awamori	phytase gene, complete cds.				
ACCESSION	102421					
VERSION	102421.1	GI:166518				
KEYWORDS	phytase.					
SOURCE	Aspergillus niger	(strain ALK0243, sub-species awamori)	DNA.			
ORGANISM	Aspergillus niger	Peizizomycolina; Eurotiomycetes;				

Db 834 AACGTGACCACTTTGATGGAATAATGCCCTTCTGAAGACATACAACTACAGCTTGGGT 893
 QY 384 gctgacgacttgactccattcgtggaacaaatggttaactctggtattatgattctac 443
 Db 894 GCAGATGACCTGACTCCTTCGGAGACAGAGCTAGTCAACTCCGGCATCAAGTCTAC 953
 QY 444 agaagatacaagggttgggtgagaagattgttccattcttagagcttctctggtctgac 503
 Db 954 CAGCGATACCAATGCTCACAAGGAACATCATTCCGTTTCATCCGATCCCTTGGCTCCAG 1013
 QY 504 agagtattgctctgctgaaagttcattgaaggtttccaactctgtaagttggctgac 563
 Db 1014 CGCTGATGCTCCCGCGAGAAATTCATGAGGGTTCACAGACCAAGCTGAAGGAT 1073
 QY 564 ccagggtgtaaccacacacaaagcttccagttattaaagttattatccagaaggtgct 623
 Db 1074 CCGTGTGCCAGCGCGGCAATCGTCCGCCAAGATCGAGTGGTGCATTTCGAGGCGCCAG 1133
 QY 624 ggttaacaaacacttgacacaggttgggttactgactgttctgaagaattctgaattgggt 683
 Db 1134 TCATCCACAACTCTCGACCCAGGACCTGCACTGTCTTTGAAGACAGCGAATTGGCC 1193
 QY 684 gacgactgaagctaaactcactgctgttttctgctccactattagagctagattgaa 743
 Db 1194 GATACCGTCGAAGCAATTTACCGCCAGGTTCCGCCCTCCTCCTCGTCAACGCTGGAG 1253
 QY 744 gctcactgacaggtgttaacttgactgacgaagcgtgttgaacttgatgacatgtg 803
 Db 1254 AAGGACCTGCTGGGGTGACTCTACAGACACAGAAAGTGAACCTACTCATGACATGTGC 1313
 QY 804 ccattgacactgttctgtagaactctgacgctactgaactgaattgtctcaattctgacttg 863
 Db 1314 TCCTTCGACACCATCTCCACGACGACCTGACACCAAGCTGTGCCGCTTCTGTGACTG 1373
 QY 864 ttcaactacagcaatgattcaataagactacttgaacttttgggttaagactactcgt 923
 Db 1374 TTCACCCATGACGATGATGATCCATGACATGACATGACATGACATGACATGACATGAC 983
 QY 984 gctagattgactactcactcactggttccagctcaaggttgggttgggttaagcaatgatt 1493
 Db 1494 CCGCGTCAACCACTGCGGCTCGGCGGACCCAGGCGCTCGGCTACGCTACGCTACGCTAC 1043
 QY 1044 aaccagctactttcccaattgaacgctacttctgacgactgaacttctcaccacacact 1553
 Db 1554 AACCCAGCTACCTTCGCGCTCACTCTACTCTCTACCGCGGACTTTTCCACGATAACGGC 1613
 QY 1104 atggtttctattttctgcttgggtttgtgacaacggtacttaagcacttcttactact 1663
 Db 1614 ATCATCTCTATCTCTTGTGCTTGGTCTGTACACGGCCTTAAGCGGCTGTCTACACG 1673
 QY 1164 tctgttgaactattgaaagaaactgacgctgactgctgcttctgactgttccactcgt 1723
 Db 1674 ACCGTGAGAAATATCACCCAGACAGATGGGTCTCTCTCTGTGACGCTTCCGTTTGT 1733
 QY 1224 gctagacttaagtgaatgatgcaatggaagctgaaagaaacattggttagagtt 1283
 Db 1734 TCAGCTGTGTAGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
 QY 1284 ttggttaacagacagagttgttccattgacggttgggttgggttgggttgggttagagtt 1343
 Db 1794 TTGGTTAATGATCGGTGTCCTCGCTGATGGGTGTCCAAATGTGATGCTTTGGGAGATGT 1853
 QY 1344 aagagagacactcgttgaaggttgggttcttctgctagatcgtggtggttaacgtggaagaa 1403
 Db 1854 ACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTTGGCGGAG 1913
 QY 1404 ttttctcgtta 1414
 Db 1914 TGTTCTGCTTA 1924

RESULT 13 ASPHYTASE

LOCUS
 DEFINITION

ACCESSION
 VERSION

KEYWORDS
 SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

exon

CDS

sig_peptide

intron

exon

mat_peptide

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Score

DB

Length

Mismatches

Indels

Gaps

ASPHYTASE 2665 bp DNA PLN 27-APR-1993
 Aspergillus niger myo-inositol hexaphosphate phosphohydrolyase gene,
 complete cds.
 M94550
 M94550.1 GI:166520
 myo-inositol hexaphosphate phosphohydrolyase; phytase.
 Aspergillus niger (library: EMBL 3; NRRL 3135) DNA.
 Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 1 (bases 1 to 2665)
 Mullanev,E.J., Gibson,D.M and Ullah,A.H.
 Positive identification of a lambda gtlI clone containing a region of
 fungal phytase gene by immunoprobe and sequence verification
 Appl. Microbiol. Biotechnol. 35, 611-614 (1991)
 92000601
 2 (bases 1 to 2665)
 Mullanev,E.J.
 Sequence of the Aspergillus niger (ficuum) phytase gene
 Unpublished (1992)

Location/Qualifiers
 1..2665
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 /tissue_lib="EMBL 3; NRRL 3135"
 <683..726,829..2188)
 join(683..726,829..2188)
 /note="phytase gene"
 /codon_start=1
 /protein_id="AAA32705.1"
 /db_xref="GI:166521"
 /translation="MGVSALLPLVLLSGVTSGLAVPASRNOSCDTVDGQYQCFSET
 SHLMOYAPFFSLANESVISEVPAGCRVTFAVLRSRGARYPTDSKGYKYSALIEEI
 QNATTFDKYAFUKYNYSLGADDTLPFGQELVNSGIFKQYVESLNRIVPFI
 SCSSVIASGKKFTGFOSTKLPDPAQPGOSPKIDVISEASSNSLTDPGCTVF
 ELSEADLVANFTAFVPSIRQRLNDLSGVTLTDTVYLMDCSFDTISTSTVD
 KLSFPCDLTHDEWINDYLSQKLYGHGAGNPLGPTQGVGYANELIARLTHSPVHD
 DTSNHTLDDSPATPLNSTLYADFSDNGILSILFALGLYNGKPLSTVVENITQT
 DGFSSATVPFASRLYVEMMCQAEPLVLRVNDRVPLHGPCVDFALGRCTRDSEV
 RGLSEARSGGDMWAECA"
 join(683..726,829..841)
 sig_peptide
 727..828
 intron
 829..2188
 exon
 842..2176
 mat_peptide
 /standard_name="phytase"
 /EC_number="3.1.3.8"
 /function="catalyzes the hydrolysis of phytic acid"
 /product="myo-inositol hexaphosphate phosphohydrolyase"

BASE COUNT 633 a 808 c 574 g 650 t

Query Match 43.4%; Score 618.2; DB 13; Length 2665;

Best Local Similarity 65.3%; Pred. No. 18e-153;

Matches 908; Conservative 0; Mismatches 483; Indels 0; Gaps 0;

QY 24 gtcgtgctactgtccattgacacacttctgctggttccacatccggtgacccgttgggtcct 83

Db 797 GTGGGACTACTGATCGCTGACTATCTGTGACAGATCACTCCGACTGGCAGTCCCGCC 856

QY 84 cgtgtaattctcactcttgtgacactgttgacggtggttaccacattgttccacagaaatt 143

Db 857 TCAGAGAAATCAATCCAGTTCGATACGGTGCATCAGGGGTATCAATGCTTCCGAGACT 916

QY 144 tctcaactgtgggttcaactctcattctcttcttcttcttcttcttcttcttcttcttcttct 203

Db 917 TCAGATCTTTGGGGTCAATACGACCGGTCTTCTCTCTCTGCGCAACCAATCGGTCTCTCC 976

QY 204 ccagacgttcccaagggtgtgtagacttcttcttcttcttcttcttcttcttcttcttcttct 263

Db	2057	TTGGTTAATGATCGCGTTGTCCTCCGCTGCATGGGTCTCCGTTGATGCTTTGGGAGATGT	2116
Qy	1344	aagagagacgaattcgttgaaggttttcttctcgtacatctggttaactggaagaa	1403
Db	2117	ACCCGGATAGCTTTCTGAGGGGGTTGAGCTTTGCTAGATCTGGGGTGATGGCGGAG	2176
Qy	1404	tggttcgctta	1414
Db	2177	TGTTTTCCTTA	2187
RESULT 14			
LOCUS	AL19451	6756 bp	DNA
DEFINITION	phytase gene locus of plasmids pAF 2-3, pAF 2-6, pAF 2-7.		PAT 10-JUN-1994
ACCESSION	AL19451		
VERSION	AL19451.1	GI:583193	
KEYWORDS	synthetic construct.		
SOURCE	artificial sequence.		
ORGANISM	1 (bases 1 to 6756)		
REFERENCE	van Gorcom, R.F.M., van Hartingsveldt, W., van Paridon, P.A.,		
AUTHORS	Veenstra, A.E., Iulien, R.G.M. and Sellen, G.C.M.		
TITLE	Cloning and expression of microbial phytase		
JOURNAL	Patent: EP 0420358-A 03-APR-1991;		
FEATURES	GIST-BROCADES N.V.		
source	Location/Qualifiers		
1..6756	/organism="synthetic construct"		
1..6756	/db_xref="taxon:32630"		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	join(210..254,357..1712)		
1..6756	/gene="phytase gene"		
1..6756	/gene="phytase gene"		
1..6756	join(210..254,357..1715)		
1..6756	/gene="phytase gene"		
1..6756	/codon_start=1		
1..6756	/transl_table=11		
1..6756	/protein_id="CAA01463.1"		
1..6756	/db_xref="GI:583194"		
1..6756	/translation="MGVSALLPLVLLSGVTSGLAVPASRNQSCDITVDGOCFSET		
1..6756	SHLWQYAPFFSLANESVISEPAGCRVTFQAQLSRHGRARYPTDSKGYKSLIET		
1..6756	QGNATVFDGKVFALKTNYSLGADLTFFQDELVNSGIRFYQYSLTRNIVPFI		
1..6756	SGSSRVIAAGKFFIEGQSTKIDPRAQPGOSPKIDVISESSNTLDPTCTV		
1..6756	EDSELADVEANFTATFVPSIKORLENDLSGLTDTETVYLMDCMSFDTISTV		
1..6756	KLSPCDLFDHDEWINDYLOSLKKYGHGAGNPLGTQGVGYANELIARLTHSPV		
1..6756	DTSSNHTLDSPTAFPLNSTLYADSFHONGIISILALGLYNGTKPLSTTVENITQ		
1..6756	DGFSSAWTVPFASRLYVENMQCQAEQELRVLRVLYNDVRVPLHGCPVDALGRCTR		
1..6756	RGLSFARSGGDWAECA"		
1..6756	/number=1		
1..6756	/gene="phytase gene"		
1..6756	/number=2		
1..6756	/gene="phytase gene"		
1..6756	/number=3		
1..6756	/gene="phytase gene"		
1..6756	/number=4		
1..6756	/gene="phytase gene"		
1..6756	/number=5		
1..6756	/gene="phytase gene"		
1..6756	/number=6		
1..6756	/gene="phytase gene"		
1..6756	/number=7		
1..6756	/gene="phytase gene"		
1..6756	/number=8		
1..6756	/gene="phytase gene"		
1..6756	/number=9		
1..6756	/gene="phytase gene"		
1..6756	/number=10		
1..6756	/gene="phytase gene"		
1..6756	/number=11		
1..6756	/gene="phytase gene"		
1..6756	/number=12		
1..6756	/gene="phytase gene"		
1..6756	/number=13		
1..6756	/gene="phytase gene"		
1..6756	/number=14		
1..6756	/gene="phytase gene"		
1..6756	/number=15		
1..6756	/gene="phytase gene"		
1..6756	/number=16		
1..6756	/gene="phytase gene"		
1..6756	/number=17		
1..6756	/gene="phytase gene"		
1..6756	/number=18		
1..6756	/gene="phytase gene"		
1..6756	/number=19		
1..6756	/gene="phytase gene"		
1..6756	/number=20		
1..6756	/gene="phytase gene"		
1..6756	/number=21		
1..6756	/gene="phytase gene"		
1..6756	/number=22		
1..6756	/gene="phytase gene"		
1..6756	/number=23		
1..6756	/gene="phytase gene"		
1..6756	/number=24		
1..6756	/gene="phytase gene"		
1..6756	/number=25		
1..6756	/gene="phytase gene"		
1..6756	/number=26		
1..6756	/gene="phytase gene"		
1..6756	/number=27		
1..6756	/gene="phytase gene"		
1..6756	/number=28		
1..6756	/gene="phytase gene"		
1..6756	/number=29		
1..6756	/gene="phytase gene"		
1..6756	/number=30		
1..6756	/gene="phytase gene"		
1..6756	/number=31		
1..6756	/gene="phytase gene"		
1..6756	/number=32		
1..6756	/gene="phytase gene"		
1..6756	/number=33		
1..6756	/gene="phytase gene"		
1..6756	/number=34		
1..6756	/gene="phytase gene"		
1..6756	/number=35		
1..6756	/gene="phytase gene"		
1..6756	/number=36		
1..6756	/gene="phytase gene"		
1..6756	/number=37		
1..6756	/gene="phytase gene"		
1..6756	/number=38		
1..6756	/gene="phytase gene"		
1..6756	/number=39		
1..6756	/gene="phytase gene"		
1..6756	/number=40		
1..6756	/gene="phytase gene"		
1..6756	/number=41		
1..6756	/gene="phytase gene"		
1..6756	/number=42		
1..6756	/gene="phytase gene"		
1..6756	/number=43		
1..6756	/gene="phytase gene"		
1..6756	/number=44		
1..6756	/gene="phytase gene"		
1..6756	/number=45		
1..6756	/gene="phytase gene"		
1..6756	/number=46		
1..6756	/gene="phytase gene"		
1..6756	/number=47		
1..6756	/gene="phytase gene"		
1..6756	/number=48		
1..6756	/gene="phytase gene"		
1..6756	/number=49		
1..6756	/gene="phytase gene"		
1..6756	/number=50		
1..6756	/gene="phytase gene"		
1..6756	/number=51		
1..6756	/gene="phytase gene"		
1..6756	/number=52		
1..6756	/gene="phytase gene"		
1..6756	/number=53		
1..6756	/gene="phytase gene"		
1..6756	/number=54		
1..6756	/gene="phytase gene"		
1..6756	/number=55		
1..6756	/gene="phytase gene"		
1..6756	/number=56		
1..6756	/gene="phytase gene"		
1..6756	/number=57		
1..6756	/gene="phytase gene"		
1..6756	/number=58		
1..6756	/gene="phytase gene"		
1..6756	/number=59		
1..6756	/gene="phytase gene"		
1..6756	/number=60		
1..6756	/gene="phytase gene"		
1..6756	/number=61		
1..6756	/gene="phytase gene"		
1..6756	/number=62		
1..6756	/gene="phytase gene"		
1..6756	/number=63		
1..6756	/gene="phytase gene"		
1..6756	/number=64		
1..6756	/gene="phytase gene"		
1..6756	/number=65		
1..6756	/gene="phytase gene"		
1..6756	/number=66		
1..6756	/gene="phytase gene"		
1..6756	/number=67		
1..6756	/gene="phytase gene"		
1..6756	/number=68		
1..6756	/gene="phytase gene"		
1..6756	/number=69		
1..6756	/gene="phytase gene"		
1..6756	/number=70		
1..6756	/gene="phytase gene"		
1..6756	/number=71		
1..6756	/gene="phytase gene"		
1..6756	/number=72		
1..6756	/gene="phytase gene"		
1..6756	/number=73		
1..6756	/gene="phytase gene"		
1..6756	/number=74		
1..6756	/gene="phytase gene"		
1..6756	/number=75		
1..6756	/gene="phytase gene"		
1..6756	/number=76		
1..6756	/gene="phytase gene"		
1..6756	/number=77		
1..6756	/gene="phytase gene"		
1..6756	/number=78		
1..6756	/gene="phytase gene"		
1..6756	/number=79		
1..6756	/gene="phytase gene"		
1..6756	/number=80		
1..6756	/gene="phytase gene"		
1..6756	/number=81		
1..6756	/gene="phytase gene"		
1..6756	/number=82		
1..6756	/gene="phytase gene"		
1..6756	/number=83		
1..6756	/gene="phytase gene"		
1..6756	/number=84		
1..6756	/gene="phytase gene"		
1..6756	/number=85		
1..6756	/gene="phytase gene"		
1..6756	/number=86		
1..6756	/gene="phytase gene"		
1..6756	/number=87		
1..6756	/gene="phytase gene"		
1..6756	/number=88		
1..6756	/gene="phytase gene"		
1..6756	/number=89		
1..6756	/gene="phytase gene"		
1..6756	/number=90		
1..6756	/gene="phytase gene"		
1..6756	/number=91		
1..6756	/gene="phytase gene"		
1..6756	/number=92		
1..6756	/gene="phytase gene"		
1..6756	/number=93		
1..6756	/gene="phytase gene"		
1..6756	/number=94		
1..6756	/gene="phytase gene"		
1..6756	/number=95		
1..6756	/gene="phytase gene"		
1..6756	/number=96		
1..6756	/gene="phytase gene"		
1..6756	/number=97		
1..6756	/gene="phytase gene"		
1..6756	/number=98		
1..6756	/gene="phytase gene"		
1..6756	/number=99		
1..6756	/gene="phytase gene"		
1..6756	/number=100		
1..6756	/gene="phytase gene"		
1..6756	/number=101		
1..6756	/gene="phytase gene"		
1..6756	/number=102		
1..6756	/gene="phytase gene"		
1..6756	/number=103		
1..6756	/gene="phytase gene"		
1..6756	/number=104		
1..6756	/gene="phytase gene"		
1..6756	/number=105		
1..6756	/gene="phytase gene"		
1..6756	/number=106		
1..6756	/gene="phytase gene"		
1..6756	/number=107		
1..6756	/gene="phytase gene"		
1..6756	/number=108		
1..6756	/gene="phytase gene"		
1..6756	/number=109		
1..6756	/gene="phytase gene"		
1..6756	/number=110		
1..6756	/gene="phytase gene"		
1..6756	/number=111		
1..6756	/gene="phytase gene"		
1..6756	/number=112		
1..6756	/gene="phytase gene"		
1..6756	/number=113		
1..6756	/gene="phytase gene"		
1..6756	/number=114		
1..6756	/gene="phytase gene"		
1..6756	/number=115		
1..6756	/gene="phytase gene"		
1..6756	/number=116		
1..6756	/gene="phytase gene"		
1..6756	/number=117		
1..6756	/gene="phytase gene"		
1..6756	/number=118		
1..6756	/gene="phytase gene"		
1..6756	/number=119		
1..6756	/gene="phytase gene"		
1..6756	/number=120		
1..6756	/gene="phytase gene"		
1..6756	/number=121		
1..6756	/gene="phytase gene"		
1..6756	/number=122		
1..6756	/gene="phytase gene"		
1..6756	/number=123		
1..6756	/gene="phytase gene"		
1..6756	/number=124		
1..6756	/gene="phytase gene"		
1..6756	/number=125		
1..6756	/gene="phytase gene"		
1..6756	/number=126		
1..6756	/gene="phytase gene"		
1..6756	/number=127		
1..6756	/gene="phytase gene"		
1..6756	/number=128		
1..6756	/gene="phytase gene"		
1..6756	/number=129		
1..6756	/gene="phytase gene"		
1..6756	/number=130		
1..6756	/gene="phytase gene"		
1..6756	/number=131		
1..6756	/gene="phytase gene"		
1..6756	/number=132		
1..6756	/gene="phytase gene"		
1..6756	/number=133		
1..6756	/gene="phytase gene"		
1..6756	/number=134		
1..6756	/gene="phytase gene"		
1..6756	/number=135		
1..6756	/gene="phytase gene"		
1..6756	/number=136		
1..6756	/gene="phytase gene"		
1..6756	/number=137		
1..6756	/gene="phytase gene"		
1..6756	/number=138		
1..6756	/gene="phytase gene"		
1..6756	/number=139		
1..6756	/gene="phytase gene"		
1..6756	/number=140		
1..6756	/gene="phytase gene"		
1..6756	/number=141		
1..6756	/gene="phytase gene"		
1..6756	/number=142		
1..6756	/gene="phytase gene"		
1..6756	/number=143		
1..6756	/gene="phytase gene"		
1..6756	/number=144		
1..6756	/gene="phytase gene"		
1..6756	/number=145		
1..6756	/gene="phytase gene"		
1..6756	/number=146		
1..6756	/gene="phyt		

Db	384	TCGAGAATCAATCCAGTTGCGATACGTCGATCAGGGGTATCAATGCTTCCGAGACT	443
Qy	144	tctcaacttggggtcaataactctcattctctcttggctgcgaatctgctattct	203
Db	444	TCGCATCTTTGGGGTCAATACACACCGTTCTCTCTGTGGCAACGAATCGGTCATCTCC	503
Qy	204	ccagaagtctccaaaggtttagagttacttctctcaagtttcttagaacacgctgt	263
Db	504	CCTGAGGTGCCCGGATGCGAGTCACTTTCGCTCAGGTCCTCTCTCCGTCATGAGGCG	563
Qy	264	agatacccaacttcttaagtcttaagaagtactctctgtcttgaagctattcaaaag	323
Db	564	CGGTATCCGACCGACTCCAAAGGCAAGAAATPACTCCGCTCTCATTTGAGGAGATCCAGCAG	623
Qy	324	aagctactgcttccaaggttaagttaagctctcttgaagctatacaactacacttgggt	383
Db	624	AAGCGACCACTTTGACGGAAATATGCGCTTCTCTGAAGACATACAACTACAGCTTGGGT	683
Qy	384	gctgacgaactgaactccattcgttgaacacaaatggttaactctggttattaagttctac	443
Db	684	GCAGATGACTGACTCCCTTCGGAGAACAGAGTAGTCAACTCCGGCATCAAGTTCTAC	743
Qy	444	agaagatacaaggtcttggcttagaagaattggttccattcgttagagctctggttctgac	503
Db	744	CAGCGGTAGAAATCGCTCACAAAGGAACATCGTTCCATTATCCGATCTCTGGCTCCAGC	803
Qy	504	agagtattgctctgctgaaagttcattgaaggttcccaattgccaactctgctaagtggctgac	563
Db	804	CGCGTATCGGCTTCGGCAAGAAATTCATCGAGGGTTCACAGGACCAACGCTGAAGGAT	863
Qy	564	ccagtgctcaaccacacaaagctctcagttatttaacggttatttccagaaggtgct	623
Db	864	CCTGTGCCACGCCGCCAATCGTCGCCCAAGATCGACGTGGTCAATTCGAGGCCAGC	923
Qy	624	ggttaacaaacacttggaccacggttctgtactgcttccgaagaactcgaattgggt	683
Db	924	TCATCCAAACAACACTTCGACCCAGGCACCTGCACGTGTTTCGAAGACAGCGAAATGGCC	983
Qy	684	gacgagctgaagctaactcactcactgctgttctgcctccacctattagagctagattgaa	743
Db	984	GATACGTCGAAGCCAAATTCACGCCAGTTCGTCCCTCCATTCGTCAACGTCGGAG	1043
Qy	744	gctcaactgccaggtgttaacttgactgacgaagacgttgttaacttgatggacatggt	803
Db	1044	AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTACCTACCTATGAGACATGTGC	1103
Qy	804	ccattcgaactgtgctgagaactctgaactcactcaattgctccactctctgacttg	863
Db	1104	TCCTTCGACACCAATCCACACAGCACCGTCGACACCAAGCTGTCCCTCTCTGTGACCTG	1163
Qy	864	ttcactcagcaaatggattcaatcacgaactactgaaattcttgggaagtactacggt	923
Db	1164	TTCAACCATGACGAATGATCAACTTACCGACTACCTTCAGTCTCTTGAAAAAGTATTCAGGC	1223
Qy	924	tacggtgctggttaacccattgggtccagctcaaggtgtggttctgtaacgaattgatt	983
Db	1224	CATGGTTCAGGTAAACCGCTCGGCCGACCCAGGCGCTCGGCTACCGTACAGAGCTCATC	1283
Qy	984	gctagattgactactctccaggttcagaacacaaactctactaacacacactttggactct	1043
Db	1284	GCCGCTGACCCACTCGCCTGTCCACGATGACACCAAGTTCCAACCACTTTGGACTCG	1343
Qy	1044	aacccagctactttcccattgaacgctacttttgactggaactctctcacgaacaact	1103
Db	1344	AGCCCGGCTACCTTTCGCTCACTCTACTCTCTACCGGACTTTTCGATGACAAACGC	1403
Qy	1104	atgggttctatttctcgtttgggtttgacaaacggtactaacgcatctgtctactact	1163
Db	1404	ATCATCTCAATCTCTTGGCTTAAAGTCTGTACACGGCAGTAAAGCCGCTATCTACCAAG	1463
Qy	1164	tctgttgaaactattgaagaactgaaggttaacgctgctcttggactgttccattgct	1223

Db	1464	ACC GTGGAATATACCCAGACAGATTCGCTCGCTGCTGGACGGTCCGTTTGGCT	1523
QY	1224	gctagagcttacgtttgaatgatgcaatgtgaagctgaaagaaacatttggttagagtt	1283
Db	1524	TGCGCTTTGACCTCGAGATGATGACAGTGTACAGCGGGAGCAGGACCGCTGCTCGGTGTC	1583
QY	1284	ttggttaacgacagagttgttccattgcacggttgtgttggttgacaacttgggttagatgt	1343
Db	1584	TTGTTTAATGATCGCGTTGTCGCCGCTGCATGGGTCTCCGGTTGATGCTTTGGGGAGATGT	1643
QY	1344	aagagagacgacttcgttgaaggtttgcttcgctagatctggttgtaactgggaaagaa	1403
Db	1644	ACCCGGGTAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGTGAATTGGCGGAG	1703
QY	1404	tggttcgctta 1414	
Db	1704	TGTTTTGCTTA 1714	
RESULT	15		
LOCUS	IL13429		
DEFINITION	IL13429	Sequence 31 from patent US 5436156.	PAT 26-JUL-1995
ACCESSION	IL13429		
VERSION	IL13429.1	GI:910770	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 6756)		
AUTHORS	Van Gorkom,R.F.M., Van Hartingsveldt,W., Van Paridon,P.A., Veenstra,A.E., Luiten,R.G.M. and Selten,G.C.M.		
TITLE	Cloning and expression of phytase from aspergillus		
JOURNAL	Patent: US 5436156-A 31 25-JUL-1995;		
FEATURES	Location/Qualifiers		
source	1..6756		
BASE COUNT	1615 a 1714 c 1712 g 1715 t		
ORIGIN	/organism="unknown"		
Query Match	43.4%;	Score 618.2;	DB 10; Length 6756;
Best Local Similarity	65.3%;	Pred. No. 1.9e-153;	
Matches 908;	Conservative	0;	Mismatches 483; Indels 0; Gaps 0;
QY	24	gtcgtgtactgtccatgtgccactgtgttggtgtccacatccggtaccgccttgggtcct	83
Db	324	GTGGGACTACTGATCGCTGACTATCTGTGCAGAGTCAGCTCCGGACTGGCAGTCCCGCC	383
QY	84	cgtggaaattctcaactctgttgacactgttgacggtgtgttaccaatgtttcccaagaatt	143
Db	384	TCGAGAAATCAATCCAGTTGGATACGGTTCGATCAGGGGTATCAATGCTTCCGAGACT	443
QY	144	tctcaactgtgggttcaataactctccattctctcttggctgcgaactctgctatttct	203
Db	444	TCGGATCTTTGGGGTCAATACGACCGCTTCTCTCTCTGGCAAACGAATCGTCACTCC	503
QY	204	ccagacgttccaaagggtgttagagttacttctgttcaagtgtgttagacacggtgtct	263
Db	504	CCTGAGGTGCGCCGGGATGCAGAGTCACTTTCGCTCAGGTCCTTCCTCGTCATGGAGCG	563
QY	264	agatacccaacttcttctaagtctaaagaagtactctgttggattgaagctattcaaaag	323
Db	564	CGGTATCCGACCGACTCCAAGGGCAAGAAATACTCCGCTCTCATTTAGGAGATCCAG	623
QY	324	aacgtactgtttcaagggttaagtcagcttcttcttggaacttacaactacactttgggt	383
Db	624	ACCCGACCACTTTGACGGAAATATFGCCTTCTCTGAAGACATACAACTACAGCTTGGGT	683
QY	384	gctgacgacttgactccacttcggtgaacacaaatggttaactctggtatgaagttctac	443
Db	684	GCAGATGACCTGACTCCCTTCGAGAACAGAGCTAGTCAACTCCGGCATCAAGTTCTAC	743

QY 444 agaagatacaaggcttggctagaagaattgttccattcgttagagcttctggttctgac 503
 Db 744 CAGCGTACGAATCGCTCAACAGGAACATGTTCCATTCATCGATCTCTGGCTCCAGC 803
 QY 504 agagtattgcttctgctgaagaattcattgaaggtttccaatctgctaaagtgtgac 563
 Db 804 CGCGTGATCGCTCCGCGAGAAATTCATGAGGGCTTCCAGAGCACCAAGCTGAAGGAT 863
 QY 564 ccaggtgtaacccacacacagcttctccagttattaaagttattattccagaaggtgct 623
 Db 864 CCTCTGCCAGCGCGGCAATTCGTCGCCCAAGATCGAGCTGCTCATTTCCGAGGCCAGC 923
 QY 624 ggttacaacaacatttggaccacggtttgttactgtctttcgaagaattctgaattgggt 683
 Db 924 TCATCCAAACAACACTCTCGACCCAGGACCTGCACGTCTTCGAGACAGCGAATTTGGCC 983
 QY 684 gacgagcttgaagcttaacttcaactgctgttttctccacactattagagctagattggaa 743
 Db 984 GATACCGTCGAAGCAATTTTCAACGCCACGTTCTGTCGCCCTCCATTCGTCAACGCTGGAG 1043
 QY 744 gctcaattgccaaggtgttaacttgaactgaagaagacgttgttaacttgatggacatgtt 803
 Db 1044 AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTGACCTACCTCATGGACATGTGC 1103
 QY 804 ccatttcgacactgttctgtagaacttctgaacttcaactcaattgtctccattctgtgacttg 863
 Db 1104 TCCTTCGACACCATCTCCACGAGCAGCGTCGACACCAAGCTGTCCCGCTTCTGTGACCTG 1163
 QY 864 ttactcaacgacgaattgattcaatacactactactgaacttcttgggttaagtactacggt 923
 Db 1164 TTCACCCATGACGAATGGATCAACTACGACTACCTTCCAGTCTTGAAGATTTACGGC 1223
 QY 924 tacggtgctggttaaccatttgggtccagctcaaggtgttgggttaacgaattgatt 983
 Db 1224 CATGGTGCAGGTAAACCGCTCGCGCCGACCCAGGCGCTCGCTACGCTAACGAGCTCATC 1283
 QY 984 gctagattgactcaactctccagttcaagaacacacacttctactaaccacacttggactct 1043
 Db 1284 GCCGCTGTGACCACTCGCTGTCCACGATGACACCAAGTTCACACCACTTTGGACTCG 1343
 QY 1044 aaccagctacttctccattgaacgctacttctgacgctacttctctcaacacacact 1103
 Db 1344 AGCCGGCTACCTTCCGCTCACTACTCTCTACGCGGACTTTTCGCATGACAAACGC 1403
 QY 1104 atggttctatttcttctgcttgggtttgtaacagcgttaacgaattgcttactact 1163
 Db 1404 ATCATCTCCATTCCTTTGCTTTAGGTCTGTACAGCGACTTAAGCCGCTATCTACCAGC 1463
 QY 1164 tctgtgaactctattgaagaactgaaggttactgcttcttggacttctccattcgt 1223
 Db 1464 ACCGTGGAGAAATACCCACAGATGGATTCGCTCTGCTTGGACGGTTCGGTTTGT 1523
 QY 1224 gctagagcttactgtgaatgtgaatgtgaagctgaaaggaacacattggttagatt 1283
 Db 1524 TCGCGTTTGTACGTCGAGATGATGACAGTGTACGCGGAGCAGGAGCGCTGCTCGTGHC 1583
 QY 1284 ttggttaacgacagagttgttccattcagcgttgggttgggtgaagaattggttagatgt 1343
 Db 1584 TTGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1643
 QY 1344 aagagagacgacttctggaaggttcttctcgttagatctgtgtggaactgggaagaa 1403
 Db 1644 ACCGGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGTGATTTGGCGGAG 1703
 QY 1404 tgttctgctta 1414
 Db 1704 TGTCTTGTCTTA 1714

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 16:39:50 ; Search time 2563.44 Seconds
(without alignments)
5258.469 Million cell updates/sec

Title: US-09-488-265-25
Perfect score: 1426
Sequence: 1 tatatgaattcatggcgctg.....tccgttaagaattcatata 1426

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

us-09-488-265-25.rst

Sat Oct 27 15:25:38 2001

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	151	10.6	802	227	AQ324539	AQ324539	mgxb0019C
2	117.8	8.3	837	226	AQ255459	AQ255459	mgxb0014M
3	113.8	8.0	581	104	AJ274007	AJ274007	AJ274007
4	88.6	6.2	753	225	AQ163004	AQ163004	mgxb0021D
C 5	81.8	5.7	699	225	AQ162040	AQ162040	mgxb0010F
6	79.4	5.6	769	227	AQ361474	AQ361474	mgxb0004I
C 7	76	5.3	810	225	AQ161556	AQ161556	mgxb0008K
C 8	72	5.0	669	227	AQ361495	AQ361495	mgxb0004K
9	45.2	3.2	1016	221	CNS04ESQ	CNS04ESQ	Tetradon
C 10	43	3.0	545	224	AQ140737	AQ140737	HS_3124_A
11	42	2.9	468	151	BF637070	BF637070	NF049F02L
12	41.8	2.9	341	109	AV054552	AV054552	AV054552
13	41.8	2.9	481	225	AQ0202816	AQ0202816	RPC111-48
14	41.2	2.9	231	25	AV129427	AV129427	AV129427
15	41.2	2.9	234	109	AV064132	AV064132	AV064132
16	41.2	2.9	236	109	AV054718	AV054718	AV054718
17	41.2	2.9	247	28	AV293643	AV293643	AV293643
18	41.2	2.9	264	108	AV002759	AV002759	AV002759
19	41.2	2.9	269	109	AV077122	AV077122	AV077122
20	41.2	2.9	278	25	AV114552	AV114552	AV114552
21	41.2	2.9	285	109	AV081956	AV081956	AV081956
22	41.2	2.9	297	109	AV060181	AV060181	AV060181
23	41.2	2.9	300	109	AV053884	AV053884	AV053884
24	41.2	2.9	306	25	AV118366	AV118366	AV118366
25	41.2	2.9	342	109	AV047696	AV047696	AV047696
26	40.4	2.8	482	30	AV423243	AV423243	AV423243
27	40.2	2.8	265	109	AV049962	AV049962	AV049962
28	40.2	2.8	295	26	AV212676	AV212676	AV212676
29	39.6	2.8	227	25	AV161297	AV161297	AV161297
30	39.6	2.8	249	109	AV056003	AV056003	AV056003
31	39.6	2.8	266	109	AV091633	AV091633	AV091633
32	39.6	2.8	268	109	AV049564	AV049564	AV049564
33	39.6	2.8	270	109	AV079904	AV079904	AV079904
34	39.6	2.8	280	109	AV078019	AV078019	AV078019
35	39.6	2.8	285	109	AV100697	AV100697	AV100697
36	39.6	2.8	289	25	AV120593	AV120593	AV120593
37	39.6	2.8	289	109	AV055109	AV055109	AV055109
38	39.6	2.8	290	109	AV053894	AV053894	AV053894
39	39.6	2.8	291	25	AV151459	AV151459	AV151459
40	39.6	2.8	296	109	AV058485	AV058485	AV058485
41	39.6	2.8	297	109	AV050799	AV050799	AV050799
42	39.6	2.8	298	109	AV067677	AV067677	AV067677
43	39.6	2.8	300	109	AV054136	AV054136	AV054136
44	39.6	2.8	303	25	AV118481	AV118481	AV118481
45	39.6	2.8	303	25	AV164971	AV164971	AV164971

ALIGNMENTS

RESULT 1
 LOCUS AQ324539/c
 DEFINITION mgxb0019C01r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0019C01r, DNA sequence.
 ACCESSION AQ324539
 VERSION AQ324539.1 GI:4116391
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 802)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL COMMENT

Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdeane@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 408.
 Location/Qualifiers
 1..802

FEATURES

source
 1..802
 /organism="Magnaporthe grisea"
 /strain="70-13"
 /db_xref="taxon:148305"
 /clone="mgxb0019C01r"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."
 BASE COUNT 166 a 218 c 209 g 208 t 1 others
 ORIGIN

Query Match 10.6%; Score 151; DB 227; Length 802;
 Best Local Similarity 55.2%; Pred. No. 5.7e-33;
 Matches 342; Conservative 0; Mismatches 266; Indels 12; Gaps 2;

QY 755 aggtgttaactgactgacgaagacgtgtgttaactgactgacatgtgacttgcattccactcac 814
 DB 614 AGGGTGAACCTCAGCGCCAAAGCGGTGAGATCATGAGCTTCTGCCGTTTGAGAC 555
 QY 815 tgtgtctagaactctgacgtactcaattgtctcattctgtgacttgcattccactcacga 874
 DB 554 CGTGGCAGATCCNCAAGCCAC-----CCTGTGCGAGTCTGCACGCTGTTTACGCAACG 501
 QY 875 cgaatgattcaatcagcactacttgcgaattcttgggttaagtactacggttgcgtgtcgtg 934
 DB 500 AGACTGGGAGGCATATGACTACTCTCCAGACACTGGGGAAGTGGTATGTTACGCAATGG 441
 QY 935 taaccattgggtccaggtccaaggtgtgtgttcgttaacgaattgattgcttagattgac 994
 DB 440 CAACCCCTGGGCTCCACCAAGGGTGGGCTTCGTCAACGAGCTCATCGCGAGCGTCT 381
 QY 995 tcactctccagttcaagacacacacttctactaacacacacttggacttaaccaggtac 1054
 DB 380 CCAAAAGCCCGTTGAAGACCCACACAAATACCAATCTGAGCTCGACACGCGCCATCGAC 321
 QY 1055 ttccattgaacgtactttgtacgtgacttctctcagcacaacactatggtttctat 1114
 DB 320 GTTCCCACTAGACAAAAAGCTGTACGCGGACTTTAGCCATGATAACGATATGCTGGCAT 261
 QY 1115 ttctctcgtttgtttgttacaacgttacttaacgaattgtctactctctgttgaatc 1174
 DB 260 CTACCCCGCGTGGGATTTACACGCCACCGCGGCCCGGATTCGGTCCCAAAAAAGGAGAG 201
 QY 1175 ta-----ttgaagaacacgtacgtgttactcgtgtcttcttggactgttccattcgtcgtag 1228
 DB 200 AAGGAGCGCGAGGAGCTACGCGGTTCTCTCGTCAGCTGGCGGTACCGTTCGAGCGAG 141
 QY 1229 agcttacgttgaatgatgcaattgtgaagctgaaaaagaaacattggttagagtttgggt 1288

Sat Oct 27 15:25:38 2001

```

Db 140 GATCTTTGTTGAAAATGACTTGGCAGGCGCAGAACGAGGAGCTTGTGAGATCTCTGGT 81
QY 1289 taacgacagagtggttccattgcacggttggttggttgacaaagttggttagatgaagag 1348
Db 80 CAACGACAGGCTGAGCCGCTGCACAACTGGCATGCCAGACAGTAGTGGTGTGTTGCACGCT 21
QY 1349 agacgacttcggtgaaggtt 1368
Db 20 GAGCAAGTTCGTTGAGAGCT 1

RESULT 2
LOCUS AQ255459 837 bp DNA GSS 23-OCT-1998
DEFINITION mgx00014M15r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgx00014M15r, DNA sequence.
ACCESSION AQ255459
VERSION AQ255459.1 GI:3779774
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 837)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 60
High quality sequence stop: 277.
High quality sequence stop: 277.

FEATURES
source
1. 837
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 248 a 154 c 188 g 243 t 4 others
ORIGIN
Query Match 8.3%; Score 117.8; DB 226; Length 837;
Best Local Similarity 56.2%; Pred. No. 2.9e-23;
Matches 221; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 748 acttgcagagtggttaactgacgaagagcaggttggttaactgtgacatgtgtccat 807
Db 140 ATTTTCAGGTGCTGGGTAAACCGCTCAAGACGCTTCTAAATCTTATGGGACGTGTCACCA 199
QY. 808 tcgacactgttgcagacttcgtgacgctactcaattgtctccattctgtactgttca 867

```

```

Db 200 TGAACACACGCGGACACTTTGAAAAAACTGGCCAACTTTCACCGCTTTCACATCTCTTTA 259
QY 868 ctacgacgaatgattacataacacactactgaactcttgggttaagtactacggttacg 927
Db 260 CGGAAGCAGATTTGGTGAATAATGTTACTTGTCCACGCGTGCAAAATATGTTACAGTACG 319
QY 928 gtgctggaacccattcgggtccagctcaaggtgtgttgcgttaacgaattgattgcta 987
Db 320 GAACGGTAATCTTTGGGCCCACTATGGCGTGGGATGGTAAACGAACTATTATGCAC 379
QY 988 gattgactcactctccagttccaagaccacactcttacttaaccacacttgaactcaacc 1047
Db 380 GATTAAACCGAAGCCAGTCGCAAGATCAACCATGTCCTCAATACGACACTTGCATGATGACC 439
QY 1048 cagctacttccattgaacgactacttctgtacgctgactctctctcagcaaacactatg 1107
Db 440 CGAAACCTTACTCTCTACAAAGCAATTTGTGTGCTGATTTATCCATACGACGATATTA 499
QY 1108 ttctatttctcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Db 500 TAGGGATTACGCGCGGCTTGGGCTATTTAAGC 532

RESULT 3
LOCUS AJ274007 581 bp mRNA EST 29-DEC-1999
DEFINITION AJ274007 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae
cDNA clone Ma1628, mRNA sequence.
ACCESSION AJ274007.1 GI:6433380
VERSION AJ274007.1
KEYWORDS EST.
SOURCE Metarhizium anisopliae.
ORGANISM Metarhizium anisopliae
Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Metarhizium.
REFERENCE 1 (bases 1 to 581)
AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.
TITLE ESR analysis of the insect pathogenic fungus Metarhizium anisopliae
JOURNAL Unpublished (1999)
COMMENT Contact: Screen SE
Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA.

FEATURES
source
1. 581
/organism="Metarhizium anisopliae"
/strain="ARSEF 2575"
/db_xref="taxon:5530"
/clone_lib="Ma1628"
/note="Vector: Unizap; Metarhizium anisopliae was grown on
insect cuticle for 24 hours. A cDNA library was
constructed in the unidirectional Lambda vector, Unizap"
BASE COUNT 138 a 153 c 146 g 144 t
ORIGIN
Query Match 8.0%; Score 113.8; DB 104; Length 581;
Best Local Similarity 58.5%; Pred. No. 3.8e-22;
Matches 237; Conservative 0; Mismatches 162; Indels 6; Gaps 2;

QY 153 tgggggtcaatactctccattcttcttggctgacgaattctgtatttccagagctt 212
Db 163 TGGGGCCACTACTCGCGGTACTCTC---AGCACACAGAGGATCTATTAGCCTTGATC 219
QY 213 ccaaaaggtgttagagttacttcttccaaagtttcttagacaggtgttagatgccca 272
Db 220 CCGTCGGGATGCGAGGTCACTTCTTCTTCAATCTTCTTCACATGGTCCCGCAACCG 279
QY 273 acttcttcaagtcttaagaagtactctgtcttggattgaagtatttcaaaagagctact 332
Db 280 ACGCCCGGCAAGTCCCAAGCCCTACAGGACCTGTTGTGAGCGATCCAAAAGACGTCAAG 339

```

```

QY 333 gctttcaagggttaagtagcttttctgaagacttaactacactttgggtgctgaacac 392
Db 340 GACTATGGGAAGGCGCTTCGAGATTTCTCAAGATTACAAGTACTACTAGTTCGACGAC 399
QY 393 tgaactcattcgttgacacaaatggttaactctctggtattgaattctacagaagatac 452
Db 400 CTCACCCCTTGGCGGAGCAGGAATGGTCAAGTCTCGAAAAAAGATTCTTCAAGCCGCTAT 459
QY 453 aaggcttgctaga---gattgtccattctgtagagcttctggttctctgacagatt 509
Db 460 CAGAAGTTAGCGGAGGAGTTCGACCCATCCCTTTGTCGAGCCCTCGGCTCTGAAGAGTC 519
QY 510 attgcttctctgtaaaagttcattgaaggtttccaatctgctaag 554
Db 520 GTCATGTGACGCGCAAGATTGTTGTCATGGCTTCTACAAAGCCCAAG 564

RESULT 4
AQ163004 753 bp DNA GSS 09-SEP-1998
LOCUS mgxb0021D19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
DEFINITION clone mgxb0021D19r, DNA sequence.
ACCESSION AQ163004
VERSION AQ163004.1 GI:3559405
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
REFERENCE 1 (bases 1 to 753)
AUTHORS Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 448.
Location/Qualifiers
1. 753
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0021D19r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 208 a 180 c 156 g 206 t 3 others
ORIGIN

Query Match 6.2%; Score 88.6; DB 225; Length 753;
Best Local Similarity 51.3%; Pred. No. 9.4e-15;
Matches 230; Conservative 0; Mismatches 212; Indels 6; Gaps 1;

```

```

QY 101 ttgtgacactgttgacggtggttaccatgtttccacagaaatttctcacttgtgggtca 160
Db 189 TTTTGATGGCGCAGCCGCGGTTTCAATTTAAAGCAGCAATCACTATATTTGGGGCA 248
QY 161 atactctcattctctcttctgtgacgaatctgctatttctccagacgttccaa---- 216
Db 249 GTATGCACCAATTTTCGTTGCACCGCTCAGCACCATCGGATTATGATTCTCTGCTACCTT 308
QY 217 --aggttgtagttagtcttctgttcaagtttctagacacggtgtagatacccaac 274
Db 309 GCCAGGGTGTAGTACTAGTGTTCGCCAAGTCTTTCAGAGGACGCGGGCTCGATACCCAAC 368
QY 275 ttcttctaagttaagaagtagtactgctgtttagtgaagctattcaaaagaacgctactgc 334
Db 369 CGCTCAAAACCGGGCGGAAATTCGACACAGATTCCACCGACTCCCAACCTCAGCTACCGG 428
QY 335 ttccaagggttaagtacgctttcttgaagaacttaacacttgggtgctgacgactt 394
Db 429 TAGTGGCGTATTGAATACTACATTAATAAACTACAAATACAAATCTCGGAGTCGAGGAATN 488
QY 395 gactccattcgttgaacacaaatggttaactctggttataagttctacagaagatacaa 454
Db 489 GAATGATTTCGGCGCGCTCGGCAACAGAGAACTCCGGGTATTNATTTTACCAAGATACCA 548
QY 455 ggcttggtagaagaagtgttccattcgttagagcttctggttctgacagagttattgc 514
Db 549 AAATCTGCCAGAGAGGACGAACCATTTATTCGCTATGATCACAANAACCGGTGTTGA 608
QY 515 ttctgctgaagaagttcattgaaggtttc 542
Db 609 TAGCCGACAACTCTGGGCTAGGGGTTTC 636

RESULT 5
AQ162040/c 699 bp DNA GSS 09-SEP-1998
LOCUS mgxb0010F06r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
DEFINITION clone mgxb0010F06r, DNA sequence.
ACCESSION AQ162040
VERSION AQ162040.1 GI:3558441
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
REFERENCE 1 (bases 1 to 699)
AUTHORS Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 285.
Location/Qualifiers
1. 699
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0010F06r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with

```

of rice world wide. It is a filamentous ascomycete with a haploid genome ($n=7$) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters

BASE COUNT	211 a	185 c	157 g	213 t	3 others
ORIGIN					

Query Match	5.6%	Score 79.4;	DB 227;	Length 769;
Best Local Similarity	51.7%;	Pred. No. 4.6e-12;		
Matches 233; Conservative	0;	Mismatches 211;	Indels 7;	Gaps 2;

[illegible]

	6
RESULT	
AQ361474	
LOCUS	DNA
DEFINITION	CUGI Rice Blast GSS
	maxb000041t2f DNA sequence.
	clocx_mayb00041t2f. DNA sequence.
	GSS
	03-FEB-1999
	Library Magnaporthe grisea genomic

Db	513	CGGTAGTGGCGGTATTGAAATCACTACATTAAAAACCTACAAATACAAATCTGGAGTCAAGGA	57
Qy	392	cttgactccattcgtgtaacacacaaatgggttaactctcgtattaaagttctcacagaagata	451
Db	573	ATTGATGATGTTTGGCGCTCCGCAACACAGAGAAC-CTCGGGTGATATTTTTTACCAAAAGATA	631
Qy	452	caaggctttggctagaagaattgttcattcgtttagagctctcgtctcgtctcgtcacagagttat	511
Db	632	CCAAAATCTGGCCGACAGGAAGCAACCATTTATTCCTATGATGACAAAAACCGCTGTT	691
Qy	512	tgcttcgtctgaaagttcattgaaggttc	542
Db	692	TGATACGCAGAAACCTTGGGCTTTAGGGTTTC	722

[illegible]

```
FEATURES
source
1. 769
Location/Qualifiers
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgx0004112f"
/tissue_type="Protoplasts"
/clone_lib="CUGI Rice Blast BAC Library"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII"
rice blast is one of the most devastating fungal diseases
```

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737

Fax: 864 656 4293
Email: rdean@clemsun.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 321.
Location/Qualifiers
1. .810

FEATURES

source
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0008K02"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request." 207 t

BASE COUNT

196 a 219 c 188 g 207 t

Query Match

Best Local Similarity 5.3%; Score 76; DB 225; Length 810;
Matches 133; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1190 cgggtacgtgcttctggactgtccattcgctgctagagcttacgttgaatgatgca 1249
Db 650 CGGTTCTGTCGACGCGGCGGTACCGTTCGACGCGAGGATGTTGTTGAAAAATGAC 591
Qy 1250 atgtgaagctgaaaagaaccattggttagagtttggtaagacagagattgttcatt 1309
Db 590 TTGCGCAGGGCAGAACGAGGAGCTGTGTGAGAACTCCTGGTCAACGACAGGGTGACGCCGT 531
Qy 1310 qcacgggtgtggtgtgacaagttgggtagatgttaagagacagacttcgtgaagttt 1369
Db 530 GCACAACTGCAGTCCGACAGTAGTGGTCTGTCACGCTGAGCAAGTTCGTGAGAGCTT 471
Qy 1370 gtcttgcgtagatctggtgttaactggggaagaatgtttcgtcttaaga 1417
Db 470 AAGCTCTGCGAGGAGTGGAGCTGCTGGGATCAATGTTTGTCTGAGA 423

RESULT 8

AQ361495/c
LOCUS 669 bp DNA GSS 03-FEB-1999
DEFINITION mgxb0004K10r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0004K10r, DNA sequence.
ACCESSION AQ361495
VERSION AQ361495.1 GI:4211334
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea

REFERENCE 1 (bases 1 to 669)
AUTHORS Yu, X., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
Clemson University
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737

Fax: 864 656 4293
Email: rdean@clemsun.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 465.
Location/Qualifiers
1. .669

FEATURES

source
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0004K10r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site.1: HindIII; Site.2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request." 233 t

BASE COUNT

185 a 119 c 131 g 233 t

Query Match

Best Local Similarity 5.0%; Score 72; DB 227; Length 669;
Matches 288; Conservative 0; Mismatches 321; Indels 14; Gaps 1;

Qy 458 ttggctagaaagattgttcattcgtagagcttgggttgcacagagttattgttc 517
Db 627 TTGCCCAAGATAAATAAATTTTATCCCTTTCTAATGGAGCGGAGTGTGTTAAAG 568
Qy 518 tgcgtgaaagttcaatgagtttccaatctgtaagtgtggtgacccaggtgtaaccc 577
Db 567 CGCTTAAATGGCTGACGGCTTTCATCAAACTTTTAAACCGATAAACTCGAATCA 508
Qy 578 acaccaagcttcacagttatttaacgttattattccagaaggtgctggttacaacac 637
Db 507 ACCGCATCAATTTCCCGCTTTCAATTTTAAATTTCTGCGAAGCAATATAACACAC 448
Qy 638 ttgggacacaggttgggtgactgcttcgaagaatcgaattgggtgacgagcttgaagc 697
Db 447 TTTAAATTAATCTCGCTTGCACCAATACGACAAATCCAAAGCTCAAAAATTTGCAAAA 388
Qy 698 t-----aacttcactgctgttttcgctccacctattagagctagattgaa 743
Db 387 ACTACCAAAAAAAGGGGAAANAATTTTGACGGAAACCGCAACCGTTGAAT 328
Qy 744 gctcactgcccaggtgttaactgactgaagaagacgttgttaactgtatgacatgtgt 803
Db 327 CAAAACCTTACCAGGGCAAAATTTATCAGTTTACGAAATTTCTCAACTTATGGAACATGT 268
Qy 804 ccattcgacactgttctagaacttctgcagctactcaattgtctccacttctgtgacttg 863
Db 267 GGTATGGAACAACACGGCTAATATTGTAAAAACGGGTATTTTATCGCCCATTTGCAATATT 208
Qy 864 ttcaactcagcagcaatggattcaatcagactacttgcactctttgttgtaagactacggt 923
Db 207 TTTGACACAAAGATTTGGGAGCGCTTTGACGGCTTTTCATATAAGTTAATGGTATACC 148
Qy 924 tacggtgctggttaacccattgggtcccagctcaagggtgtgtgttcgttaacgaattgatt 983
Db 147 CACTTTTATGGTAACCCCTTTAGGGCCGTCATTGGGAGTAGGTTGGGTAATGAACITATC 88
Qy 984 gctagattgactcactctccagttcgaagacacacacttctactaacacatttggactct 1043
Db 87 GCGAGGTTATTACGAAGACCTGTACAAGATTTCCACCAGCACCACCAACCAATTAAGTCTGC 28


```

clone 1810031C08, mRNA sequence.
AV054552
VERSION AV054552.1 GI:5154299
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 341)
AUTHORS Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T., Hara
,A., Hayatsu,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M., Kawai,J.,
Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsuma,H., Oda,H., Owa,C.,
Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara
,Y., Suzuki,H., Suzuki,H., Tateo,M., Tomaru,Y., Tominaga,N.,
Watanabe,S., Yagame,M., Yamamura,T., Yokota,T., Yoshino,M.,
Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs
JOURNAL Unpublished (1999)
COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermostactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
            Location/Qualifiers
                1..341
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="1810031C08"
                    /clone_lib="Mus musculus pancreas C57BL/6J adult"
                    /sex="male"
                    /tissue_type="pancreas"
                    /dev_stage="adult"
BASE COUNT      86 a      60 c      101 g      94 t
ORIGIN
Query Match      2.9%; Score 41.8; DB 109; Length 341;
Best Local Similarity 52.6%; Pred.No.0.34;
Matches 91; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 1166 tgttgatctattgaagaactgacgctgtacgctgtcttcttggaactgttcaccatcgctgc 1225
      || || || || || || || || || || || || || || || || || || || || ||
Db 151 TGTGGGCGCTGCCAGCAGCTGAGTGGCGCTGCTGGTGTGTGTGTTCACCTCCTGC 210
      || || || || || || || || || || || || || || || || || || || || ||
QY 1226 tagagcttcattgaatgatgcaatgtgaagctgaaaggaaacccattggtagagtttt 1285
      || || || || || || || || || || || || || || || || || || || || ||
Db 211 CGCGCGCCCGCTGTGAGGAGAAGAAAGTGGAGGCAAGAAAGAAATCCGAGGAGATTGA 270
      || || || || || || || || || || || || || || || || || || || || ||
QY 1286 ggtaacgacacagagtgttccattcacggtgtgtgtgtgtgtgacaaagtgggta 1338
      || || || || || || || || || || || || || || || || || || || || ||
Db 271 AGATGACATGGGCTTTGTTTTTTTTCACTAACCTGTTTTTGTAAAGTTGGCTA 323
      || || || || || || || || || || || || || || || || || || || || ||

RESULT 13
AQ020816
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
human.

```

RESULT	13
AQ02816	
LOCUS	481 bp DNA
DEFINITION	RPC11-48J14.TK RPCI-11 Homo sapiens genomic clone RPCI-11-48J14, DNA sequence.
ACCESSION	AQ02816
VERSION	AQ02816
KEYWORDS	AQ02816.1 GI:3613607 GSS. human.
SOURCE	

AUTHORS

JOURNAL,

CONCLUSION

RIKEN
3-1-1 YOUNDAI TOWN

Tel: 81-298-36-9145

Email: genome-res@rt

trehalose and its application for

Location/Qualifiers

```
/db xref="taxon:1009
```

```
/clone lib="Mus muscul
```

```
/sex= male :
/tissue type="small
```

```

55 3 15 0 72 2
55 3 15 0 72 2
55 3 15 0 72 2

```

Best Local

0; Indels 83; Mismatches 0; conservative 0; mismatches 0;

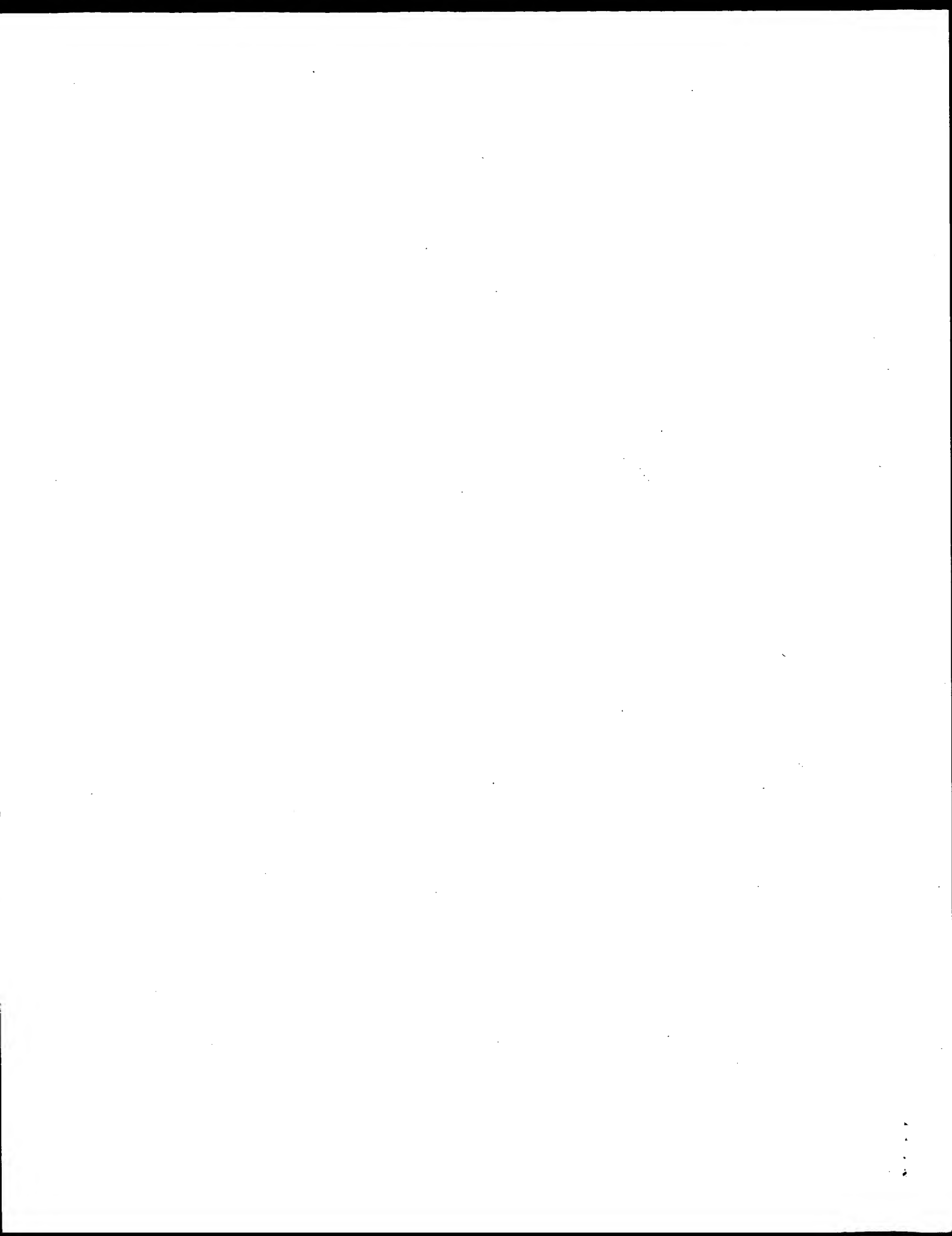
[illegible][illegible]

Oy 1325 c t a g e o c t + n e s e r i e + u p + m e d + l e a +

[illegible][illegible]

Search completed: October 20, 2001, 18:33:00
Job time: 6790 sec

[illegible]



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 13:24:54 ; Search time 48.78 Seconds
(without alignments)
729.264 Million cell updates/sec

Title: US-09-488-265-26

Perfect score: 2470

Sequence: 1 MGVFVLLSIATLFGSTSGT.....DFEGLSPARSGNNEEFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1866	75.5	467	1 JN0889	3-phytase (EC 3.1.1.1)
2	1862	75.4	467	1 JN0656	3-phytase (EC 3.1.1.1)
3	1801	72.9	441	1 JN0482	3-phytase (EC 3.1.1.1)
4	403	16.3	467	1 PABYC	acid phosphatase (EC 3.1.1.1)
5	399	16.2	467	2 S53476	acid phosphatase (EC 3.1.1.1)
6	399	16.2	467	2 S48996	acid phosphatase (EC 3.1.1.1)
7	394	16.0	467	1 PABYC	acid phosphatase (EC 3.1.1.1)
8	370	15.0	468	2 S52495	acid phosphatase (EC 3.1.1.1)
9	350.5	14.2	479	1 JN0715	3-phytase (EC 3.1.1.1)
10	350	14.2	468	2 JC4285	acid phosphatase (EC 3.1.1.1)
11	346.5	14.0	479	1 JN0890	acid phosphatase (EC 3.1.1.1)
12	326	13.2	453	1 A25326	acid phosphatase (EC 3.1.1.1)
13	301	12.2	463	2 S14119	acid phosphatase (EC 3.1.1.1)
14	301	12.2	463	2 T39929	thiamin-repressibl
15	156	6.3	465	2 JEO369	histidine acid pho
16	154.5	6.3	468	2 A86233	hypothetical prote
17	134	5.4	755	2 T19118	acid phosphatase h
18	130	5.3	449	2 T15933	acid phosphatase h
19	122.5	5.0	423	2 A33395	acid phosphatase (
20	119	4.8	380	2 T16883	hypothetical prote
21	117.5	4.8	421	2 S14742	hypothetical prote
22	112.5	4.6	423	1 S06167	acid phosphatase (
23	108.5	4.4	381	2 JH0152	acid phosphatase (
24	108	4.4	354	2 T21241	acid phosphatase (
25	107.5	4.4	602	2 A36715	hypothetical prote
26	106.5	4.3	730	2 T16455	exo-poly-alpha-gal
27	104.5	4.2	413	2 JY0087	hypothetical prote
28	104.5	4.2	413	2 B85636	glucose-1-phosphat
29	103.5	4.2	496	2 T44987	periplasmic glucos
					aldehyde dehydroge

ALIGNMENTS

RESULT 1

JN0889

3-phytase (EC 3.1.3.8) A precursor - Aspergillus awamori

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase: phyA protein

C:Species: Aspergillus awamori

C>Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C/Accession: JN0889

R:Pidington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0889

A:Molecule type: DNA

A:Residues: 1-467 <PID>

A:Cross-references: GB:L02421; NID:gl66518; PIDN:AAAL6898.1; PID:gl66519

A:Experimental source: strain ALK0243

A:Note: part of the sequence, including the amino end of the mature protein, was conf

C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic

C:Genetics:

A:Gene: phyA

A:Introns: 15/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine: phosphoprotein; ph

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-467/Product: 3-phytase A #status experimental <MAT>

F:27,59,105,120,207,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #

F:81,361/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match

Best Local Similarity 75.5%; Score 1866; DB 1; Length 467;

Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTSGTALPGNSHCDVDGGYQCFPELSHLWGQYSPFFSLADE 60

Db 1 MGVSALLPLLYLAGVTSGLVAPASRNSQCTDVGQYQCFSETSLWGQYAPFFSLANE 60

QY 61 SAISPDVPGKRVTFVQVLSRHGARYPTSSKKYSALTEAIOKNATAFKGYAFUKTYN 120

Db 61 SAISPDVPGKRVTFVQVLSRHGARYPTSSKKYSALTEAIOKNATAFKGYAFUKTYN 120

QY 121 YILGADLLTPFGEQMVNSGKIFRYRYKALARKIVFVRASGSDRVIAAEKFIQFQSA 180

Db 121 YILGADLLTPFGEQMVNSGKIFRYRYKALARKIVFVRASGSDRVIAAEKFIQFQSA 180

QY 181 KIADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPTR 240

Db 181 KIADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPTR 240

QY 241 ARLEAHLPGVNIITDEEVNLMDCPFDTVARTSDATQLSPPCDFLTHDEWIOYDYLQSLG 300

Db 241 ARLEAHLPGVNIITDEEVNLMDCPFDTVARTSDATQLSPPCDFLTHDEWIOYDYLQSLG 300

30 100.5 4.1 693 1 S61067
31 100.5 4.1 1225 2 T16346
32 100 4.0 438 2 S64682
33 100 4.0 452 2 T20556
34 98.5 4.0 642 2 E69144
35 98 4.0 386 1 JH0610
36 98 4.0 844 2 T37690
37 97.5 3.9 630 2 S77346
38 97 3.9 1350 2 G36793
39 96.5 3.9 374 2 C75105
40 96 3.9 537 2 S54770
41 96 3.9 721 2 T27570
42 96 3.9 888 2 T46726
43 96 3.9 1433 2 A71444
44 95.5 3.9 433 2 S74407
45 95.5 3.9 609 1 KSPSCY

homoaconitate hydr
hypothetical prote
acid phosphatase (
hypothetical prote
probable formate C
acid phosphatase (
hypothetical prote
hypothetical prote
hypothetical prote
secreted acid phos
secreted acid phos
probable LTR retro
homoserine dehydro
copper resistance

Db 241 ORLENDLSGVLTDTDEVYLYMDMSFDTISTSTVDTKLSPFCDLTFHDEWHLYDYLQSLK 300
Qy 301 KYYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHHTLDSNPATFPLNATLYADFS 360
Db 301 KYYGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADFS 360
Qy 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCAEKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAMTVPFASKRLXYEMMQCAEQEP 420
Qy 421 LVRVLNDRVVPVPLHGGVCKLGRKRDDEFVEGLSPARSGGNWEECEFA 467
Db 421 LVRVLNDRVVPVPLHGGVCKLGRKRDDEFVRLGLSFARSGGDWAECSA 467
RESULT 2
JN0656
3-phytase (EC 3.1.3.8) A precursor - Aspergillus niger
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
C:Species: Aspergillus niger
C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0656; S28456
R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Hartevelde, G.M.; Gouka, R.J.; Suykerbuyk, M.
A.M.J.J.
Gene 127, 87-94, 1993
A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)
A:Reference number: JN0656; MUID:9322284
A:Accession: JN0656
A:Molecule type: DNA
A:Residues: 1-467 <VAN>
A:Cross-references: GB:Z16414; NID:q2392; PIDN:CAA78904.1; PID:g2393
A:Experimental source: strain NREL3135
A:Note: parts of the sequence, including the amino end of the mature protein, were confirmed
C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and inorganic orthophosphate from phyt
C:Genetics:
A:Gene: phyA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
F:1-23/Domain: signal sequence #status experimental <SIG>
F:24-467/Product: 3-phytase A #status experimental <WAT>
F:27,59,105,120,207,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #sta
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 75.4%; Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 1.7e-141;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
Qy 1 MGVFVLLSTATLFGSTGALGRGNSHCDTVDGGYQCFPEISHLWQGYSPFFSLADE 60
Db 1 MGVSALLPLXLSGVTSLGLAVPASRQSCDITVDQGYQCFSETSHLWQYAPFFSLANE 60
Qy 61 SAISPDVPGKRVTFVQVLSRHGARYPTSSKSKYKYSALIEIAOKNATFAKGYAFKLTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTDSKSKYKYSALIEIQONATTFDQGYAFKLTYN 120
Qy 121 YTLGADDTLPGEQOMNSGKIFVRRYKALARKTVPPVRSAGSDRVTAIAEKIEGQPSA 180
Db 121 YSLGADDTLPGEQELNSGKIFQRESTRNTRVPIRSGSGSRVTAIAEKIEGQFSQ 180
Qy 181 KLADPGANPHQASPVNVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPP 240
Db 181 KLKOPRAQPGQSKIDVWISSESSNNTLDPGCTVFEDELADTVEANFTAVFAPP 240
Qy 241 ARLEAHLPGVNLTDVVDVMDMCFDTVARTSDATQLSPECDFTHDEWIDYDYLQSLG 300
Db 241 ORLENDLSGVLTDTDEVYLYMDMSFDTISTSTVDTKLSPFCDLTFHDEWHLYDYLQSLK 300
Qy 301 KYYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHHTLDSNPATFPLNATLYADFS 360
Db 301 KYYGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADFS 360

Qy 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCAEKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAMTVPFASKRLXYEMMQCAEQEP 420
Qy 421 LVRVLNDRVVPVPLHGGVCKLGRKRDDEFVEGLSPARSGGNWEECEFA 467
Db 421 LVRVLNDRVVPVPLHGGVCKLGRKRDDEFVRLGLSFARSGGDWAECSA 467
RESULT 3
JN0482
3-phytase (EC 3.1.3.8) A - Aspergillus ficum
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein
C:Species: Aspergillus ficum
C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
C:Accession: JN0482; PN0023
R:Ullah, A.H.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 192, 747-753, 1993
A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemic
A:Reference number: JN0482; MUID:93249451
A:Accession: JN0482
A:Molecule type: protein
A:Residues: 1-441 <ULL>
A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 178, 45-53, 1991
A:Title: Cyclohexanediol modification of arginine at the active site of Aspergillus
A:Reference number: PN0023; MUID:91298982
A:Accession: PN0023
A:Molecule type: protein
A:Residues: 48-70 <ULL>
C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydrate (Asn) (covalent) #sta
F:58,338/Active site: Arg, His #status predicted
F:59/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 72.9%; Score 1801; DB 1; Length 441;
Best Local Similarity 76.0%; Pred. No. 1.3e-136;
Matches 333; Conservative 42; Mismatches 63; Indels 0; Gaps 0;
Qy 27 NSHSCDTVDQGYQCFPEISHLWQGYSPFFSLADESAISPOVPGKRCRVTFVQVLSRHGARY 86
Db 4 NOSSCDTVDQGYQCFSETSHLWQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 63
Qy 87 PTSSKSKYKYSALIEIAOKNATFAKGYAFKLTYNITLGADDTLPFGQOMVNSGKIFYRR 146
Db 64 PDSKSKYKYSALIEIQONATTFDQGYAFKLTYNISLGADDTLPFGQELVNSGKIFYQR 123
Qy 147 YKALARKTVPPVRSAGSDRVTAIAEKIEGQFSQAKLADPGANPHQASPVNVIPEGAGY 206
Db 124 YESLTRNVPPIRSGSGSRVTAIAEKIEGQFSQTKLKDPAQPGQSSPKIDVWISSESS 183
Qy 207 NNTLDHGLCTAFEESELGDDVEANFTAVFAPPFARLEAHLPGVNLTDVVDVMDMCP 266
Db 184 NNTLDPGCTVFEDELADTVEANFTAVFVPSIRORLENDLSGVTLTDTTEVYLYMDMCSF 243
Qy 267 DTAVRTSDATQLSPECDFTHDEWIDYDYLQSLGKYYGYGAGNPLGPAQGVGFVNELIAR 326
Db 244 DTISTSTVDTKLSPFCDLTFHDEWIDYDYLQSLKYYGAGNPLGPTQGVGYANELIAR 303
Qy 327 LPHSPVQDHTSTNHHTLDSNPATFPLNATLYADFSHNDTMVSIFFALGLYNGTKPLSTTSV 386
Db 304 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHNDGLIISILFALGLYNGTKPLSTTV 363
Qy 387 ESTEETDGYAASWTVPFAARAYVEMMQCAEKEPLVRVLNDRVVPVPLHGGVCKLGRCKR 446
Db 364 ENITQDGFSSAMTVPFASKRLXYEMMQCAEQEPPLVRVLNDRVVPVPLHGGVCKLGRCTR 423
Qy 447 DDFVEGLSPARSGGNWEE 464

A:Reference number: S53458

A:Accession: S53476

A:Molecule type: DNA

A:Residues: 1-467 <BUS>

A:Cross-references: EMBL:L28920; NID:g1616966; PIDN:AAC09508.1; PID:g456155; MIPS:YAR071

R:Chen, J.Y.; Gong, Y.; Ao, S.Z.

Acta Biochim. Biophys. Sin. 21, 437-444, 1989

A:Title: The primary structure of acid phosphatase gene PHO11 in *S. cerevisiae* and comparison with other yeast acid phosphatase genes

A:Reference number: JCI018

A:Accession: JCI018

A:Molecule type: DNA

A:Residues: 1-16, 'L', 18-149, 'H', 151-353, 'O', 355-422, 'G', 424-467 <CHE>

A:Note: This paper is in Chinese, with an English abstract

C:Genetics:

A:Gene: SGD:PHO11

A:Cross-references: SGD:S0000094; MIPS:YAR071w

A:Map position: 1R

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-467/Product: acid phosphatase #status predicted <MAT>

F:74/Active site: Arg #status predicted

F:75/Active site: His (phosphohistidine intermediate) #status predicted

F:97,162,192,250,315,356,439,445,461/Binding site: carbohydrate (Asn) (covalent) #status

Query Match

Best Local Similarity 16.2%; Score 399; DB 2; Length 467;

Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGCGYQCFPEISHLMQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHG 83

DB 21 PLGKLSIDIKIGTQTETFPFL---GGSPYYSFPGDYGISRDLPSCMKQVMVGRHG 76

QY 84 ARPTSSKSKKYSALLIATOKNATKGYAFK-----TNYNTL-----GADDL 128

DB 77 ERYPTVSKAKSIMTWTWYKLSNYTGQFSGALSFNDDEFFIRDTKMLETTILANSVNL 136

QY 129 TPE-GQQMVNSGKIKYRYKALARKIVP-VRASGSDRVIAAEKFIIEGFSQAKLADPG 186

DB 137 NPYTGEMNAKRHARDLAQGYGVVENQTSFAVTSNSNRCHDTAQYFIDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241

DB 191 -----FNISLQITISEASAGANTSAHSHCPAWDD-DVNDILKDKYTKYLSGIK 240

QY 242 RLEAHLPGVNLTDVVDVLMDCPDTVARTSDATQSPCDLFTDHEWTOYDYLQSLGK 301

DB 241 RLNKENKGLNLTSSDANTFFAWCAYENAR-----GYSDICNIFTKDELVRFSYQDLET 295

QY 302 YGYGAGNPLGPAQGVGVNLIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSE 361

DB 296 YYOTGPGYDVRSVGANLNASVKLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFFGALYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEPL 421

DB 338 DTDILNLTITIGIDDKNNLTAEHVPFMENT--FHRSWYVPGQARVYTEKFOC--SNDTY 393

QY 422 VRVLNDRVPLHGGCGVDKLGCRKDDF-----VEGLSFAR 457

DB 394 VRYVINDAVVPIETCTGPGFSCSEINDFYAEKRVAGTDFLK 436

RESULT 6

S48996

N:Alternate names: PHO12 - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999

A:Accession: S48996; S59659

R:Macri, C.

submitted to the EMBL Data Library, February 1994

A:Description: The sequence of *S. cerevisiae* cosmid 9177.

A:Reference number: S46671

A:Accession: S48996

A:Molecule type: DNA

A:Residues: 1-467 <MAC>

A:Cross-references: EMBL:U00029; NID:g551322; PIDN:AAB69729.1; PID:g458917; MIPS:YHR2

R:Xu, L.

submitted to the EMBL Data Library, January 1995

A:Reference number: S59658

A:Accession: S59659

A:Molecule type: DNA

A:Residues: 1-16, 'L', 18-81, 'AR', 84-149, 'H', 151-467 <XUL>

A:Cross-references: EMBL:U19789; NID:g847754; PIDN:AAA73479.1; PID:g847755

C:Genetics:

A:Gene: SGD:PHO12

A:Cross-references: SGD:S0001258; MIPS:YHR215w

A:Map position: 8R

C:Superfamily: yeast acid phosphatase

C:Keywords: phosphoric monoester hydrolase

Query Match

Best Local Similarity 16.2%; Score 399; DB 2; Length 467;

Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGCGYQCFPEISHLMQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHG 83

DB 21 PLGKLSIDIKIGTQTETFPFL---GGSPYYSFPGDYGISRDLPSCMKQVMVGRHG 76

QY 84 ARPTSSKSKKYSALLIATOKNATKGYAFK-----TNYNTL-----GADDL 128

DB 77 ERYPTVSKAKSIMTWTWYKLSNYTGQFSGALSFNDDEFFIRDTKMLETTILANSVNL 136

QY 129 TPE-GQQMVNSGKIKYRYKALARKIVP-VRASGSDRVIAAEKFIIEGFSQAKLADPG 186

DB 137 NPYTGEMNAKRHARDLAQGYGVVENQTSFAVTSNSNRCHDTAQYFIDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241

DB 191 -----FNISLQITISEASAGANTSAHSHCPAWDD-DVNDILKDKYTKYLSGIK 240

QY 242 RLEAHLPGVNLTDVVDVLMDCPDTVARTSDATQSPCDLFTDHEWTOYDYLQSLGK 301

DB 241 RLNKENKGLNLTSSDANTFFAWCAYENAR-----GYSDICNIFTKDELVRFSYQDLET 295

QY 302 YGYGAGNPLGPAQGVGVNLIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSE 361

DB 296 YYOTGPGYDVRSVGANLNASVKLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFFGALYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAKEPL 421

DB 338 DTDILNLTITIGIDDKNNLTAEHVPFMENT--FHRSWYVPGQARVYTEKFOC--SNDTY 393

QY 422 VRVLNDRVPLHGGCGVDKLGCRKDDF-----VEGLSFAR 457

DB 394 VRYVINDAVVPIETCTGPGFSCSEINDFYAEKRVAGTDFLK 436

RESULT 7

PABVCC

N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c

C:Species: Saccharomyces cerevisiae

C:Date: 30-Sep-1991 #sequence_revision 09-Sep-1994 #text_change 05-Nov-1999

A:Accession: S48259; S45960; S05794; A25241; S44674

R:Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.

Yeast 10, 1363-1381, 1994

A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.

A:Reference number: S48255; MUID:95208357

A:Accession: S48259

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-467 <MAN>

A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55597.1; PID:g476050

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
 R:Feidmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45927
 A:Accession: S45960
 A:Molecule type: DNA
 A:Residues: 1-467 <FE2>
 A:Cross-references: EMBL:Z35961; NID:G536362; PIDN:CAA85045.1; PID:G536363; GSPDB:GN0000
 Nucleic Acids Res. 12, 7721-7739, 1984
 R:Bayja, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
 A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeas
 A:Reference number: S05794; MUID:85037940
 A:Accession: S05794
 A:Molecule type: DNA
 A:Residues: 1-218, 'MKT', 222-467 <BAJ1>
 A:Cross-references: EMBL:X01080; NID:G4148; PIDN:CAA23557.1; PID:G758281
 A:Note: The authors translated the codon AAT for residue 134 as Asp and TAC for residue
 R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G
 Mol. Cell. Biol. 6, 1855-1865, 1986
 A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
 A:Reference number: A93074; MUID:87064474
 A:Accession: A25241
 A:Molecule type: DNA
 A:Residues: 1-44 <TAI>
 C:Genetics:
 A:Gene: SGP:PHO3; MIPS:YBR092C
 A:Cross-references: SGD:S0000296; MIPS:YBR092C
 A:Map position: 2R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoric monoester hydroly
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-467/Product: acid phosphatase, constitutive #status predicted <MAT>
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (coval
 F:337/Active site: His #status predicted

Query Match 16.0%; Score 394; DB 1; Length 467;
 Best Local Similarity 27.0%; Pred. No. 9.7e-24;
 Matches 117; Conservative 62; Mismatches 191; Indels 64; Gaps 13;
 QY 49 GOYSPFFSLADESAISDPVKGCRVTFOVLSRHGARYPTSSKSKYSALIEAIQKNATA 108
 Db 42 GGAGPVFSPGCGISRDLEGGCEMQLQMLARHGERIPYPSYKSLNTRYQ 101
 QY 109 FGKGYAFLEK-TYNYTLGADP-----LTPF-GEQQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFLNDYDEFFIRDDDLMEETTFSNDVNLNPTTGEMDAKRHAREFLAQGYMFE 161
 QY 153 KIVPF-VRASGSDRVITASAEKFELEGFQSAKLADPGANPHQASPVINVIPEGAGYNTLD 211
 Db 162 NQTSFFIFFAASSERVHDTAQYFIDGL-----GDQFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFEESELGDDVEANFTAVFAPPIRARLEAHLPGVNLTDVQVNLMDCPEDTVA 270
 Db 211 AGNACPCWDE-DANDILDKYDITYLDDLAKEKLNKENKGLNLSKDNLTFAWCAYELNA 269
 QY 271 RYSDATQSPFCDFLTHDEWIQDYDYLQSLGKYGYGAGNPLGPAQGVFNVELIARLTHS 330
 Db 270 R-----GYSDVCIDFTEDELRYSGQDLVSFYQDQPGYDMIRSVGANLFNATLKLKOS 324
 QY 331 PVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFFALGLYNGTGPLSTTSVESIE 390
 Db 325 ETQD-----LKVLSFTHTDITLNYLTAGIIDDKNLTAAYVPPFG 366
 QY 391 ETDGAASTVTPFAARAYVEMMOCAEKEPLVRLVNDVWVPLHGGVDBKLGCRKDDP- 449
 Db 367 NT--FHKSNIYVPOGARVYTEKFQC--SNDTYVRYVINDAVVETCTSGPGFSCEINDEY 422
 QY 450 -----VEGLSFAR 457
 Db 423 DYAEARVAGTDLK 436

RESULT 8

S52495
 acid phosphatase homolog YDL024c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D2815
 C:Species: Saccharomyces cerevisiae
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
 C:Accession: S52495; S67556
 R:Andre, B.; Vissers, S.; Urrestarazu, L.
 submitted to the EMBL Data Library, February 1995
 A:Description: The sequence of a 42 kb segment located on the left arm of chromosome
 A:Reference number: S52492
 A:Accession: S52495
 A:Molecule type: DNA
 A:Residues: 1-468 <AND>
 A:Cross-references: EMBL:Z48432; NID:G683669; PIDN:CAA88335.1; PID:G683673
 A:Experimental source: strain S288C
 R:Urrestarazu, L.A.; Andre, B.; Vissers, S.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67335
 A:Accession: S67556
 A:Molecule type: DNA
 A:Residues: 1-468 <URR>
 A:Cross-references: EMBL:Z74072; NID:G1430996; PIDN:CAA98583.1; PID:G1430997; MIPS:YD
 A:Experimental source: strain S288C
 C:Genetics:
 A:Map position: 4L
 C:Superfamily: yeast acid phosphatase

Query Match 15.0%; Score 370; DB 2; Length 468;
 Best Local Similarity 25.1%; Pred. No. 8.2e-22;
 Matches 116; Conservative 62; Mismatches 174; Indels 110; Gaps 14;

QY 49 GOYSPFFSLADESAISDPVKGCRVTFOVLSRHGARYPTSSKSKYSALIEAIQKNATA 108
 Db 43 GGSAPYFSPFANYGIPDIDPEGCHLTQVMIGRGEYPTRSEAKDIFEVWYKISNTYTKG 102
 QY 109 FGKGYAFLEK-TYNYTL-----GADLTTPF-GEQQMVNSGKIFRYRYKALAR 152
 Db 103 YEGSLSFNLNGYEFFIPDESLEMETTLQNSIDVNLNPTGEMNAKRHAREFLAKYKIME 162
 QY 153 KIVPF-VRASGSDRVITASAEKFELEGFQSAKLADPGANPHQASPVINVIPEGAGYNTLD 211
 Db 163 NCTNFPIFTNSKRIYDTAQYFAEL-----GDGFNISLQ 197
 QY 212 HGCTAFEESELG-----DDVEANFTAVFAPPIRARLEAHLPGVNL 253
 Db 198 -----TLSENSSSGANTLAAKSSCPNNSNANDILMSYSDYLENLSDRLENKGLNLS 253
 QY 254 DEDVNLMDCPEDTVARISDATQLSPFCDFLTHDEWIQDYDYLQSLGKYGYGAGNPLGP 313
 Db 254 RKDAALFSCAPEL-----NAKYSNICDIFSAAELIHYSETDLTSTFYQNGPGYKLIK 308
 QY 314 AQGVGVN---ELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTWYSIFF 370
 Db 309 SIGANLFNATVKLIROSAH-----LDOKVWLSFTHTDITLNYLT 347
 QY 371 ALGLYNGTGPLSTTSVESIEETDGYAASVTVPFAARAYVEMMOCAEKEPLVRLVNDV 430
 Db 348 TAGLIDDTNLNTNHPFRDHS--YHRSWVPOGARVYTEKFQC--SNDSYRYVYVNDV 403
 QY 431 VPLHGGVDBKLGCRKDDFE-----GLSFARSGNNEEC 465
 Db 404 VPIESCSSGPGFSCEGTTEYAYAKDLRGVSE-----YEDC 439

RESULT 9

JN0715
 3-phytase (EC 3.1.3.8) B precursor - Aspergillus ficum
 N:Alternate names: pH 2.5-optimum acid phosphatase
 C:Species: Aspergillus ficum
 C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

Sat Oct 27 15:25:40 2001

Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast *Pichia pastoris*: Characterization
A:Reference number: JC4285; MUID:96001238
A:Accession: JC4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:g881955; PIDN:AAA85503.1; PID:g881956
A:Experimental source: GS115
C:Genetics:
A:Gene: pho1
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:8/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:345/Active site: His #status predicted

Query Match 14.2%; Score 350; DB 2; Length 468;
Best Local Similarity 25.4%; Pred. No. 3.3e-20;
Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;

QY 6 VLLSIATL---FGSTGALPGRNHSCDFTVDGGYOCFFEISHLMGOYSPF--FSLADE 60
Db 10 IILALATLQSVFAVELQHLVGLVNDPRPQRT-DDQYNILRHGLG-GPYLYGNGWIAAE 67
QY 61 SAISPDVPKGRVTFVOVLSRHGARYPTSSKSKYSALIE-AIQKNATAFKGYAFKTY 119
Db 68 SEI-----ESTIQAHLLMRHGERYSTNVGKLEALYKLLDADVEVPTGPLSPFDY 122
QY 120 NYTLG---ADLLTPFGFQOMVNSGIRP---YRYKAL-----ARKIVFFVRASGDR 165
Db 123 DYFVSDAAWYEQETTKGFYSGLTAFDFTLRERYDHLINTSEEGKL--SVWAGSQER 180
QY 166 VIASKEFIEGFSQAKLADPGANPHQASPVINVIPE---GAGYNNLTLDHGLCT----A 217
Db 181 VDTAKYFAOGFMKSNYTD-----WVEVVALEEEKSOGSLNLTARISCPNYNSHI 230
QY 218 FEESLGDGV---EANTFAVFAPPFIRAEHLPGVNLTDVYVNLMDMCPDFVARTSD 274
Db 231 YKGDFFNDIAERAD-----RLNTLSPGNITADDIPTIALYCGFELNVRGE- 278
QY 275 ATQLSPFDLTFHDEWIOYDYLQSLGKYYGACGNPLGPAOGVGFVNELIARLTHSPVD 334
Db 279 ----SSFCVLSREALTYAVLRDLGYNVNGNPNLGTIGYVIAN----- 321
QY 335 HTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPALLGYNKTK-PLSTTSVESIEETD 393
Db 322 --ATQLLENTAD-PRDYPLYFSFSDTDLLOVFTSLGLFNVDLPL-----DQIQFOT 373
QY 394 GYAASWTVPFAARAYVEMMQC--EAEKEPLVRVLVNDVRVPLHGGVCKLGRKRDDEVE 451
Db 374 SFKSTEIVPMGARLLTERLLCTVEGEKYYVYRILNDAAVFLPLSDCSGPGFSCPLNDYVS 433
QY 452 GLSFARSGGNWEE 464
Db 434 RLEALNEDSDFAE 446

RESULT 11
JN0890
acid phosphatase (EC 3.1.3.2) precursor - *Aspergillus awamori*
C:Species: *Aspergillus awamori*
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0890
R:Pidington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0890
A:Molecule type: DNA
A:Residues: 1-479 <PID>

Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast *Pichia pastoris*: Characterization
A:Reference number: JC4285; MUID:96001238
A:Accession: JC4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:g881955; PIDN:AAA85503.1; PID:g881956
A:Experimental source: GS115
C:Genetics:
A:Gene: pho1
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:8/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:345/Active site: His #status predicted

Query Match 14.2%; Score 350.5; DB 1; Length 479;
Best Local Similarity 27.9%; Pred. No. 3.1e-20;
Matches 117; Conservative 53; Mismatches 159; Indels 91; Gaps 16;

QY 49 GOYSPFSLADESAISPDVPKGRVTFVOVLSRHGARYPTSSKSKY-SALIEAIOKNAT 107
Db 53 GPYSERVY---GIADPTGCEVDQVIMVKGHGERYSPSAGKSIIEALAKVYSINT 108
QY 108 AFKGYAFKTYNYTL-----GADLLT-PF-GEQOMVNSGIRFYRYKAL--ARKIVFF 157
Db 109 EYKGDALFNDWTYVYVNECYNAETTSYAGLLDAYNHGNDYKARYGHLWNGETVWVPF 168
QY 158 VRASGSDRVIASAEKLEGFQSAKLADPGANPHQASPVINVIPEGAGYNNLTLDHGLCTA 217
Db 169 F-SSGIRGVETARKFGEF-----FGYNYSTNAALNII 201
QY 218 FEESLGLG-----DDVEANFTAVFAPPFIRAEHLPGVNLTDVYVNLMD 262
Db 202 SESEVMGADSLTPCTDNDQTCNLTQYQLPQFKVAARLNSONPGMNLTSADVYNLMV 261
QY 263 MCPEDTVAATSDATQLSPFDLTFHDEWIOYDYLQSLGKYYGACGNPLGPAOGVGFVNE 322
Db 262 MASEFELNAR-----PFSNWINAFTQDEWVSFGYVEDLNYCYGPGDKNMAAVGAVTANA 316
QY 323 LIARLTHSPVDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPALLGCLY--NGTKP 380
Db 317 SLTLNQGP-----KEAGP-----LFFNLADHTNITPILALGVLPNEDLP 358
QY 381 LSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEA-----EKEPLVRVLVNDVRVPLHGC 436
Db 359 L-----DRVAFGNPYSIGNIVPMGGHLLTIERLSQATALSADKGYVRLVNLNEAVLPFND 413

RESULT 10
JC4285
acid phosphatase (EC 3.1.3.2) precursor - yeast (*Pichia pastoris*)
N:Alternate names: orthophosphoric monoester phosphohydrolase (acid optimum); Pho
C:Species: *Pichia pastoris*
C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999
C:Accession: JC4285
R:Payne, W.E.; Gannon, P.M.; Kaiser, C.A.

A:Cross-references: GB:L02420; NID:g166481; PIDN:AAAL6897.1; PID:g166482
 A:Experimental source: strain ALK0243
 C:Comment: The highly similar enzyme from A. ficum has been shown to have 3-phytase (EC 3.1.3.2)
 C:Genetics:
 A:Gene: aph

A:Introns: 261/1; 300/2; 335/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phospho
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-479/Product: 3-phytase #status predicted <NAT>
 F:81.337/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted
 F:106,191,227,250,315,340,425,442,458/Binding site: carboxylate (Asn) (covalent) #status

Query Match 14.08; Score 346.5; DB 1; Length 479;
 Best Local Similarity 27.48; Pred. No. 6.5e-20;
 Matches 115; Conservative 54; Mismatches 160; Indels 91; Gaps 15;

QY 49 GOYSPFFSLADESAISPDVPGKCRVTFQVLSRRGARYPTSSKSKY-SALIEAIQKAT 107

Db 53 GPYSERSVY-----GIARDPPTSCVDQVIMVKRGERYPSPSAGKDEALAKVISINT 108

QY 108 AFKCKYAFKTYNTL-----GADDT-PP-GEQMVNSGKIFRYKAL--ARKIVPF 157

Db 109 EYKGLAFLDWTYYVNECYNAETTSQYAGLLDAYNHGNDYKARYGHLWNGETVVPF 168

QY 158 VRASGRVIAASAEKFTGQSAKLADPGANPHQASPVINIIPEGAGYNTLDHGLCTA 217

Db 169 F-SSGGRVITARKFEGP-----FGYNTSTNAALNII 201

QY 218 FESELG-----DDVEANFTAVFAPPTR---ARLEAHLPGVNLTDDEVNLM 262

Db 202 SESEVMGADSLTPTCDTNDQTTCDNLTYQLPQKVAARLNQPCGNLTASDVYLMV 261

QY 263 MCPEDTARTSDATQSPFCDLTFHDEWIOYDYLSGLKGYGAGNPLGPAQGVGVNE 322

Db 262 MASFELNAR-----PFSNWINAFQDEWISFGYVEDLUNYICAGPGDKNMAAVCAVYANA 316

QY 323 LIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLY--NGTKP 380

Db 317 SLTLNAGPKP-----AGSLFFNFADHTNTPTLALGVLIPNEDLP 358

QY 381 LSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEA-----EKEPLRVLVNDRVPLHGC 436

Db 359 L-----DRVAFGNPYSIGNIVPMGGHLTIERLSCQATLSDEGTYVRLVNLNEAVLPENDC 413

RESULT 12

A25326

acid phosphatase (EC 3.1.3.2) precursor [similarity] - fission yeast (Schizosaccharomyces
 C:Species: Schizosaccharomyces pombe
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: A25326; T50405
 R:Elliot, S.; Chang, C.; Schweingruber, M.E.; Schaller, J.; Rickli, E.E.; Carbon, J.
 J. Biol. Chem. 261, 2936-2941, 1986
 A:Title: Isolation and characterization of the structural gene for secreted acid phosphatase
 A:Reference number: A25326; MUID:86140050
 A:Accession: A25326

A:Molecule type: DNA

A:Residues: 1-453 <ELL>

A:Cross-references: GB:M11857; NID:g173422; PIDN:AAA5321.1; PID:g173423

R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

A:Experimental source: strain 972h(-); clone p1 p4G3

A:Reference number: Z25068

A:Accession: T50405

A:Molecule type: DNA

A:Residues: 1-453 <RIB>

A:Cross-references: EMBL:AL137099; PIDN:CAB68657.1; GSPDB:GN00067; SPDB:SPBP4G3.02

A:Experimental source: strain 972h(-); clone p1 p4G3

C:Genetics:

A:Gene: phol; SPDB:SPBP4G3.02

A:Map position: 2

C:Superfamily: yeast acid phosphatase
 C:Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
 F:68/Active site: Arg #status predicted
 F:69/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 13.2%; Score 326; DB 1; Length 453;
 Best Local Similarity 24.8%; Pred. No. 2.6e-18;
 Matches 103; Conservative 63; Mismatches 201; Indels 48; Gaps 8;

QY 35 DGGYOCFPEISHLWGQYSPFFSLADESAISPDVPGKCRVTFQVLSRRGARYPTSSKSKK 94

Db 27 DGKDFKHELTSRSPYHKPYF-----YGPSIDFPTTCKIKQVHTLQRHGSRNPTGGNAF 81

QY 95 YSALIEAIQK-----NATAFKGYAFLKTYNTL---GADDTLTFEGEQMVNSGK 142

Db 82 DAVGLANFQORLLNGSVDPVTSVGNPLSFVPTPTVTEAANADALSSSGRVELFDMGRQ 141

QY 143 FYRKYALARKIVFPVRASGSDRVIAAEKFTGQSAKLADPGANPHQASPVINIIPE 202

Db 142 FYERYHELFASTYNIYTAQQRVVDSALWYGYGMF-----CEDVINFTNYILVSENA 194

QY 203 GAGYNTLDHGLCTAFESSELGDDVEANFTAVFAPPTRARLEAHLPGVNLTDDEVNLM 262

Db 195 TAGNSLSSYNACPASDADFTTPALEARNVYMPPIRQLNPFYSNYNLTDNDILNLYG 254

QY 263 MCPEDTARTSDATQSPFCDLTFHDEWIOYDYLSGLKGYGAGNPLGPAQGVGVNE 322

Db 255 ICSVEIALQ-----DYSEFCKLFSVDFLNEFYEGDLSFSYGMGNSVKWGSIFGAVANS 309

QY 323 LIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKPLS 382

Db 310 LANSL--RSVENNTQ-----QVFFAETHDANIIPVETALGFFTDNTPEN 351

QY 383 TTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEEKPEPLRVLVNDRVPLHGC 437

Db 352 PLPTSQVSHSKASEFVFAGNLITELFCQEDSKY-YVRHLVNEEVFPLSDCG 405

RESULT 13

S14119

acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
 C:Accession: S14119; T40455
 R:Yang, J.; Schweingruber, M.E.
 Curr. Genet. 18, 269-272, 1990
 A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schiz
 A:Reference number: S14119; MUID:91064763
 A:Accession: S14119

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-463 <YAN>

A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA40258.1; PID:g5007

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z21931

A:Accession: T40455

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: DNA

A:Residues: 1-463 <LYN>

A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN00067; SPDB:SPBC428.03c

A:Experimental source: strain 972h-; cosmid c428

C:Genetics:

A:Gene: SPBC428.03c

A:Map position: 2

C:Superfamily: yeast acid phosphatase

C:Keywords: phosphoric monoester hydrolase

Query Match

Best Local Similarity 12.2%; Score 301; DB 2; Length 463;

Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

Db 269 AIKDH-----SDFCSTFTSEFLNFEYDSDLQOAYGGGPVSEWASTGLGAYINNLADEL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFALGLYNGTKPLSTTSVES 388
 Db 323 -----RNVNTPDFDRK-----VFLAFTHDSNIIPIVEAALGFFPDITFONPLPTDK 367
 QY 389 IEETDGYAASWTVPFAARAYVEMMOCEAEKPLRVLVNDVVRVPLHGGCGVDKLGK- - - - C 444
 Db 368 NIYYSOKTSFVFPFAGNLITELFFC-SDSKYYVRHLVNOQVYPLIDCGYGPSTSDGLC 426
 QY 445 KRDDFVEGLSFARSGGN 461
 Db 427 ELQAYLNSPIRANSTSN 443

RESULT 15
 JE0369
 histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
 C:Accession: JE0369
 R:Mullane, E.J.; Ullah, A.H.J.
 Biochem. Biophys. Res. Commun. 251, 252-255, 1998
 A:Title: Identification of a histidine acid phosphatase (phya)-like gene in Arabidops
 A:Reference number: JE0369; MUID:99009256
 A:Accession: JE0369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-465 <MUL>
 A:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 6.3% Score 156; DB 2: Length 465;
 Best Local Similarity 22.2%; Pred. No. 0.00012;
 Matches 100; Conservative 65; Mismatches 200; Indels 86; Gaps 20;

QY 35 DGGYOCPEISHLMGOVSPFFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTSSSKK 94
 Db 16 DGGFDVRHULSTV-TRYSTSKDVTQNLIEGNSVPSECTPIHLNIVARHGTRSPKTKLRE 74
 QY 95 YSAL--TEATOKNATPK-----GKAFIKTVNTLGGADLTTPFGEQOMVNSGI 141
 Db 75 MESLAGREFKELVRDAEARKLSKIPKMGQWK--SPWEGKVGKGLIRQGEDELYQLGI 132
 QY 142 KFYRYKALARK-----IVPFVRASGSDRVIASAEKFIEGFOSAK-LADPGANPHQASPIV 196
 Db 133 RYRERFSLFEEDYHPDVYVITRATQIPRASASAVAFMGLESEKGNLPGGRNRAFA---- 188
 QY 197 NVIIPEGAGYNTLHGLCTAFEESEELGDDVEANFTAVFAPPIRARLEAHLPGVNLTD 256
 Db 189 --VTSENRASTDKLRFEC-----QNY-----KSYRKAKPEAVDKLKEP 226
 QY 257 VVNLMDMC--PFDTVARTSDATQSLSPFCDLTHDEWTDYQYDLSLKGKYYGYGAGNPLGP 313
 Db 227 VLNKITASVAKRYDILKFKQDILSSLWFLCKQVALLW-----TDDLEVLFLKGYGNSLNY 281
 QY 314 AGQGVFNVELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFFA-L 372
 Db 282 KMGVPLLEDVL-----HSMEEAIKAREKLP--PGSYE-KARL--RFAHAETIVPPFSCLL 332
 QY 373 GLYNGTKPLSTTSVESIE-----ETDGYAASWTVPFAARAYVEMMOCEAEK 420
 Db 333 GLF-----LDGSEFEKIQEKEPLELPQPPKTRDFRGSTMAPFGGNLILVLYSCPAESP 387
 QY 421 --LVRYLVNDVVRVPLHGGCGVDKLGKRDFF 449
 Db 388 KYEVQVHLNHEPIAVPGC--DGKDFCPLDEF 416

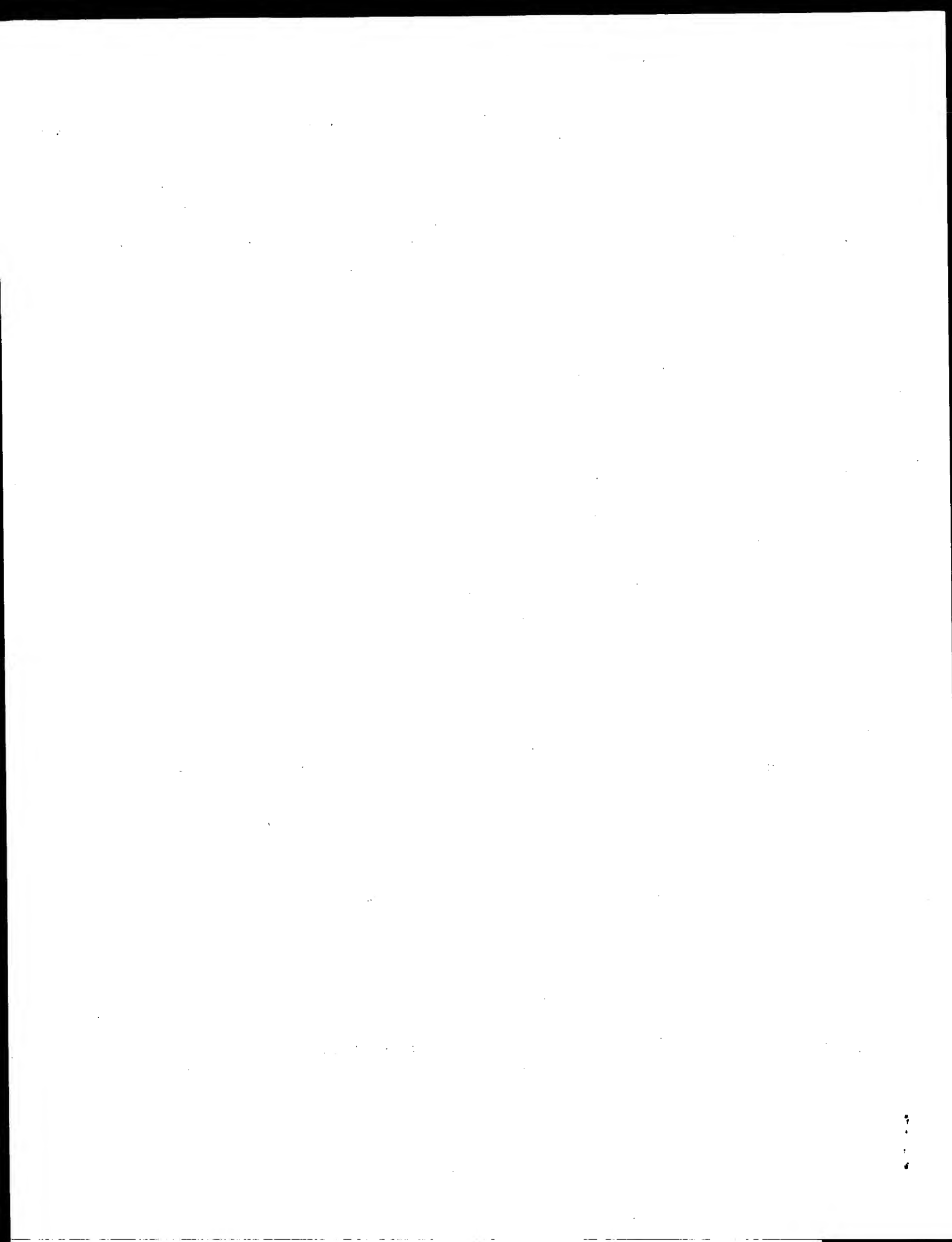
Search completed: October 26, 2001, 15:10:27
 Job time: 6333 sec

QY 51 YSPFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTSS-----SKSKYSALIE----- 100
 Db 43 HEPYF-----NGPTTISFESCAIKQVHLLQHRGSRNPtGDDTATDVSSAQYIDIFQKLL 97
 QY 101 --AIQKNATAFKGYAFUKTYNYTL---GADDLTPFGEQOMVNSGIKRYRYKALARKIV 155
 Db 98 NGSIPVNFSEYPENPLFYFKHMTFVKAENADOLSSSGRIELFDLGRQVFEYRYELFDYDV 157
 QY 156 PFVRASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPE--GAGYNTLHDG 213
 Db 158 YDINTAQERVVDSAEWFSYGM-----FGDMQKTNF-----IVLPEDDSAGANSLSMYN 208
 QY 214 LCTAFEESELGDD-VEANFTA---VFAPPPIRARLEAHL-PGVNLTDDEVNLMDCPFD 268
 Db 209 SCPVYEDNNDIDENTEAHTSRNFKLPIANRLNKYFDGSGYNTLTVSDVRSLYYICVYEI 268
 QY 269 VARTSDATQSLSPFCDLTHDEWTDYQYDLSLKGKYYGYGAGNPLGPAQGVGFVNELIARLT 328
 Db 269 ALRDN-----SDFCSTFTSEFLNFEYDSDLQOAYGGGPVSEWASTGLGAYINNLADEL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADSHDNTMVSIFALGLYNGTKPLSTTSVES 388
 Db 323 -----RKGVNNSDRK-----VFLAFTHDSNIIPIVEAALGFFPDITPEHPLPTDK 367
 QY 389 IEETDGYAASWTVPFAARAYVEMMOCEAEKPLRVLVNDVVRVPLHGGCGVDKLG 442
 Db 368 NIYYSOKTSFVFPFAGNLITELFFC-SDSKYYVRHLVNOQVYPLIDCGYGPSTSDGLC 426

RESULT 14
 T39929
 thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39929
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.
 submitted to the EMBL data Library, May 1998
 A:Reference number: Z21857
 A:Accession: T39929
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-463 <LYN>
 A:Cross-references: EMBL:AL023286; PIDN:CAAL1863.1; GSPDB:GN00067; SPDB:SPBC21H7.03c
 A:Experimental source: strain 972h; cosmid c21H7
 C:Genetics:
 A:Gene: SPDB:SPBC21H7.03c
 A:Map position: 2
 C:Superfamily: yeast acid phosphatase

Query Match 12.2% Score 301; DB 2: Length 463;
 Best Local Similarity 24.3%; Pred. No. 2.7e-16;
 Matches 106; Conservative 71; Mismatches 198; Indels 62; Gaps 13;

QY 51 YSPFSLADESAISDPVPGKGRVTFVQVLSRHGARYPTS--SKSKYSA-----L 98
 Db 43 HEPYFDGLDSA-----FPTEICIQVHLLQHRGSRNPtGDDTATDVSSQYLNQFQKLL 97
 QY 99 IEAIQKNATAFKGYAFUKTYNYTL---GADDLTPFGEQOMVNSGIKRYRYKALARKIV 155
 Db 98 NGSIPVNFSEYPENPLCFIKQWTPVIDAENADOLSSSGRIELFDLGRQVFEYRYELFDYDV 157
 QY 156 PFVRASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPEG--AGYNTLHDG 213
 Db 158 YDINTAQERVVDSAEWFSYGLFGDKMFE-----KTNFILLSGKAAGANSLSMYN 208
 QY 214 LCTAFEESEL-----GDDVEANFTAVFAPPPIRARLEAHL-PGVNLTDDEVNLMDCPFD 268
 Db 209 ACVPFKDNFNHKNATDAHAHVNRNIFIEPIVNRKAKYFDSSYKLTINDVRSFLYICEYEI 268
 QY 269 VARTSDATQSLSPFCDLTHDEWTDYQYDLSLKGKYYGYGAGNPLGPAQGVGFVNELIARLT 328



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:34 ; Search time 29.88 Seconds
(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGVEVLLSIATLFGSTSGT.....DFVEGLSFARSGNNEECPA 467

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1866	75.5	467	1	PHYA_ASPAW
2	1862	75.4	467	1	PHYA_ASPNG
3	1813.5	73.4	463	1	PHYB_EMENI
4	403	16.3	467	1	PPA5_YEAST
5	399	16.2	467	1	PPAB_YEAST
6	399	16.2	467	1	PPAC_YEAST
7	394	16.0	467	1	PPA3_YEAST
8	370	15.0	468	1	PPAD_YEAST
9	352.5	14.3	479	1	PHYB_ASPNG
10	350	14.2	468	1	PPAL_PICPA
11	346.5	14.0	479	1	PHYB_ASPAW
12	340.5	13.8	469	1	PPA5_KLULA
13	326	13.2	453	1	PPA1_SCHPO
14	301	12.2	463	1	PPA2_SCHPO
15	134	5.4	755	1	PPAX_CAEEL
16	122.5	5.0	423	1	PPAL_RAT
17	117.5	4.8	421	1	PPAL_MOUSE
18	112.5	4.6	423	1	PPAL_HUMAN
19	108.5	4.4	381	1	PPAP_RAT
20	107.5	4.4	602	1	PERX_ERWCH
21	104.5	4.2	413	1	AGP_ECOLI
22	100.5	4.1	693	1	LYS4_YEAST
23	98	4.0	386	1	PPAP_HUMAN
24	98	4.0	844	1	YD25_SCHPO
25	97	3.9	1350	1	VG72_HSVII
26	96	3.9	721	1	NICA_CAEEL
27	95.5	3.9	433	1	DHOM_SVNY3
28	95.5	3.9	609	1	COPA_PSESM
29	95	3.8	938	1	TRG1_ECOLI
30	95	3.8	3587	1	TYCB_BACBR
31	94	3.8	411	1	PPAY_CAEEL
32	94	3.8	467	1	ATPX_BACFI
33	94	3.8	1527	1	MRP3_HUMAN
34	93.5	3.8	396	1	ARSL_AQUAE
35	93.5	3.8	646	1	IRG1_MOUSE
36	93.5	3.8	2485	1	PTND_HUMAN
37	91.5	3.7	526	1	CLOS_CLOHT
38	91.5	3.7	586	1	VP40_ILTVT
39	91	3.7	413	1	PPAW_CAEEL
40	90	3.6	366	1	ODPB_YEAST
41	90	3.6	428	1	GSA_STAAU
42	90	3.6	626	1	BGAL_LEULA
43	89.5	3.6	845	1	CC47_YEAST
44	89	3.6	417	1	AGP_PRORE
45	89	3.6	1191	1	DNBI_MCMVS

ALIGNMENTS

RESULT 1		PHYA_ASPAW		STANDARD;		PRT: 467 AA.	
ID	PHYA_ASPAW						
AC	P34753;						
DT	01-FEB-1994 (Rel. 28, Created)						
DT	01-FEB-1994 (Rel. 28, Last sequence update)						
DT	01-OCT-2000 (Rel. 40, Last annotation update)						
DE	3-OXYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE						
DE	3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE						
DE	PHOSPHOHYDROLASE A).						
GN	PHYA OR PHY.						
OS	Aspergillus awamori.						
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;						
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.						
OX	NCBI_TaxID=105351;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=ALK0243;						
RA	MEDLINE=94040796; PubMed=8224894;						
RA	Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,						
RA	Miettinen-Oinonen A., Nevalainen H., Rambossek J.A.;						
RT	"The cloning and sequencing of the genes encoding phytase (phy) and						
RT	pH 2.5 optimum acid phosphatase (aph) from Aspergillus niger var.						
RT	awamori."						
RL	Gene 133:55-62(1993).						
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE						
CC	FROM PHYTASE.						
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =						
CC	D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.						
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use by non-profit institutions as long as its content is in no way						
CC	modified and this statement is not removed. Usage by and for commercial						
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/						
CC	or send an email to license@isb-sib.ch).						
CC	-----						
DR	EMBL; L02421; AAA16898.1; -						
DR	PIR; JN0889; JN0889.						
DR	HSSP; P34752; LIHP.						
DR	InterPro; IPR000560; -						
DR	Pfam; PF00328; acid_phosphat; 1.						
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.						
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.						
KW	Hydrolase; Glycoprotein; Signal.						
FT	SIGNAL 1 23						
FT	CHAIN 24 467						
FT	ACT_SITE 82 82						
FT	ACT_SITE 361 361						
FT	DISULFID 31 40						
FT	DISULFID 71 414						
FT	DISULFID 215 465						
FT	DISULFID 264 282						
FT	DISULFID 436 444						

O66908 aquifex aso
P54987 mus musculus
Q12923 homo sapien
P09870 clostridium
P23984 infectious
Q09451 caenorhabdi
P32473 saccharomyc
Q04092 staphylococ
Q02603 leuconostoc
P38132 saccharomyc
Q52309 providencia
P30672 murine cyto

```

FT CARBOHYD 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 51075 MW; 118E28A5D7EC661 CRC64;

Query Match 75.5%; Score 1866; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 3.5e-143;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVSFVLLSIALFGSTGALPGRNHSDCTVGGVQCQFPEISHLWGQYSPFFSLADE 60
Db 1 MGVSFVLLSIALFGSTGALPGRNHSDCTVGGVQCQFPEISHLWGQYSPFFSLANE 60
QY 61 SATSPDPKGRVTFVQVLSRHGARYPTSSKSKYSALTEALQKNATAPKGYAFKLTYN 120
Db 61 SATSPDPKGRVTFVQVLSRHGARYPTSSKSKYSALTEALQKNATAPKGYAFKLTYN 120
QY 121 YTLGADLTPFGQOMVNGIKFYRYKALARKIVFVRASGSDRVASAEKFIQFQSA 180
Db 121 YTLGADLTPFGQOMVNGIKFYRYKALARKIVFVRASGSDRVASAEKFIQFQSA 180
QY 181 KLDADPGANPHQASPVNLTIPGAGVNTLIDGLCTAFPESELGDDVEANFTAFAPPT 240
Db 181 KLDADPGANPHQASPVNLTIPGAGVNTLIDGLCTAFPESELGDDVEANFTAFAPPT 240
QY 241 ARLEALHPGVNLTDEVDVNLMDMCPEDTWTARTSDATQSLPFCDFLTHDEWIOYDYLQSLG 300
Db 241 ARLEALHPGVNLTDEVDVNLMDMCPEDTWTARTSDATQSLPFCDFLTHDEWIOYDYLQSLG 300
QY 301 KYGYGAGNPLGPAQGVNFVNLIELARTHSVPQVODHTNTHLDSNPATPLNATLYADFS 360
Db 301 KYGYGAGNPLGPAQGVNFVNLIELARTHSVPQVODHTNTHLDSNPATPLNATLYADFS 360
QY 361 HDNFWISFFALGLNGTKPLSTSVESIEDTGYAASWTVPFAARAYVEMMOCEAEKEP 420
Db 361 HDNFWISFFALGLNGTKPLSTSVESIEDTGYAASWTVPFAARAYVEMMOCEAEKEP 420
QY 421 LVRVLVNDVRVPLHGGCVKLGRCRKRDDFVEGLSFRASGNGWEECEFA 467
Db 421 LVRVLVNDVRVPLHGGCVKLGRCRKRDDFVEGLSFRASGNGWEECEFA 467

```

```

RESULT 2
PHYA_ASPPG STANDARD; PRT; 467 AA.
AC P34752;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93252284; PubMed=8387447;
RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
RA Gouka R.J., Suykerbuyk M.E.G., Luiten R.G.M., van Paridon P.A.,
RA Selten G.C.M., Veenstra A.E., van Gorcom R.F.M.,
RA van den Hondel C.A.M.J.J.;

```

```

RT "cloning, characterization and overexpression of the phytase-encoding
RT gene (phyA) of Aspergillus niger.";
RL Gene 127:87-94(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mullanev E.J.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-464.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93249451; PubMed=8387289;
RA Ullah A.H.J., Dischinger H.C. Jr.;
RT "Aspergillus ficuum phytase: complete primary structure elucidation
RT by chemical sequencing.";
RL Biochem. Biophys. Res. Commun. 192:747-753(1993).
RN [4]
RP SEQUENCE OF 71-93.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=91298982; PubMed=1648914;
RA Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
RT "Cyclohexanedione modification of arginine at the active site of
RT Aspergillus ficuum phytase.";
RL Biochem. Biophys. Res. Commun. 178:45-53(1991).
RN [5]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=89160685; PubMed=2852807;
RA Ullah A.H.J.;
RT "Aspergillus ficuum phytase: partial primary structure, substrate
RT selectivity, and kinetic characterization.";
RL Prep. Biochem. 18:459-471(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=97307250; PubMed=9164457;
RA Kostrewa D., Gruninger-Leitch F., D'Arcy A., Broger C., Mitchell D.,
RA van Loon A.P.;
RT "Crystal structure of phytase from Aspergillus ficuum at 2.5-A
RT resolution.";
RL Nat. Struct. Biol. 4:185-190(1997).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTASE.
CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -1- BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN
CC FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
CC USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER
CC THE NAME PHYTASE NOVO.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z16414; CAA78904.1; -.
DR EMBL: M94550; AAA32705.1; -.
DR PIR: JN0482; JN0482.
DR PIR: JN0023; JN0023.
DR PIR: JN0656; JN0656.
DR PDB: 1IHP; 18-MAR-98.
DR InterPro: IPR000560; -.
DR Pfam: PF00328; acid.phosphat. 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 467 3-PHYTASE A.
FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).

```

van Loon A.P.;
 "Cloning of the phytases from Emericella nidulans and the
 thermophilic fungus Talaromyces thermophilus";
 Biochim. Biophys. Acta 1353:217-223(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U59803; AAB96871.1; ..
 DR InterPro; IPR000560; ..
 DR Pfam; PF00328; acid_phosphat_1;
 DR PROSITE; PS00617; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00718; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 POTENTIAL.
 FT CHAIN 20 463 3-PHYTASE B.
 FT ACT_SITE 80 80 REQUIRED FOR BINDING SUBSTRATE
 FT (BY SIMILARITY).
 FT ACT_SITE 81 81 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 357 357 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 463 AA; 51786 MW; ECC5827DIE1C82A2 CRC64;

QY 421 LVRVLNDRVPLHGGVDKLGCRKDDFVEGLSFARSGGNWEECF 466
 Db 416 LVRVLNDRVPLHGGVDKLGCRKDDFVEGLSFARSGGNWEECF 461

RESULT 4
 PPAS_YEAST STANDARD; PRT; 467 AA.
 ID PPAS_YEAST
 AC P00635;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE REPRESENTABLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
 GN PHO5 OR YBR093C OR YBR0814
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 18-26.
 RX MEDLINE=83168913; PubMed=6300772;
 RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
 RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
 of repressible acid phosphatase contains a signal peptide.";
 RL Nucleic Acids Res. 11:1657-1672(1983).
 RN [2]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 18-45.
 RX MEDLINE=85037940; PubMed=6093051;
 RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RT "Structural analysis of the two tandemly repeated acid phosphatase
 genes in yeast.";
 RL Nucleic Acids Res. 12:7721-7739(1984).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R.; Ehle S., Vetter I., Feldmann H.;
 RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 RN [4]
 RP SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE=87064474; PubMed=3537710;
 RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
 RT Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
 RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
 cluster within the acid phosphatase multigene family of Saccharomyces
 cerevisiae.";
 RL Mol. Cell. Biol. 6:1855-1865(1986).
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
 CC -1- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; V01320; CAA24630.1;
 CC EMBL; X01079; CAA25555.1;
 CC EMBL; X78993; CAA55598.1;
 CC EMBL; Z35962; CAA85046.1;
 CC EMBL; X01080; CAA25556.1;
 CC PIR; S05795; PABYC.
 CC PIR; B25241; B25241.
 CC HSP; P34752; LIHP.
 CC SGD; S0000297; PHO5.

DR InterPro: IPR000560;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1
 FT CHAIN 18 467 REPRESSIBLE ACID PHOSPHATASE.
 FT ACT_SITE 75 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT CARBOHYD 97 97 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 36 36 D -> Y (IN REF. 1).
 FT CONFLICT 130 130 H -> G (IN REF. 1).
 FT CONFLICT 294 294 H -> Q (IN REF. 1).
 FT CONFLICT 446 446 S -> V (IN REF. 1).
 FT CONFLICT 462 463 AS -> DT (IN REF. 1).
 FT CONFLICT 466 466 R -> K (IN REF. 1).
 SQ SEQUENCE 467 AA; 52858 MW; DC3G9504BC2D3D0C CRC64;

Query Match 16.3%; Score 403; DB 1; Length 467;

Best Local Similarity 26.2%; Pred. No. 4.8e-25;
 Matches 126; Conservative 61; Mismatches 216; Indels 78; Gaps 15;

QY 6 VLLSTATLFGSTALPGRNHSCDTVGQYQCFPEISHLWGQYSPFFSLADESAISP 65
 Db 5 VVYSLAASLANAGTI--PLGKLADVDKIGTKQDIFPL---GGAGPYVSPFGDYGISR 58
 QY 66 DVPKGRVTVQVLSRRGARYPTSSKKYSALIEATOKNATAPKGYAEFLK-TYNTLGL 124
 Db 59 DLPECEMKQLQMVGRHGERYPTVSLAKTIKSTWYKLSNTYTRFNGSLFLNDDYEFFIR 118
 QY 125 ADD-----LTPF-GEQMVNSGKIFRYRKALKARKIVPF-VRASGSDRVIA 168
 Db 119 DDDDDLEMTTTFANSDVDVLPYTGEMNAKRHARDELAQYGTAVENQTSFAVTSNKRCHD 178
 QY 169 SAEKRIEFGSAKLADPGANPHQASPVINII-----PEGAGYNNITLDHGLCTAFESEEL 223
 Db 179 TQAYFIDG-----LGDQ-----FNITLQTVSEASAGANTLSACNSCPAW-DYDA 222
 QY 224 GDDVEANFTAVFAPPFIRARLEAHLPGVNLTDDEVDVNLMDMCPDFTVARTSDATOLSPFCD 283
 Db 223 NDDIVNEYDTTVDLDDIAKRLNKENKGLNLTSTDASTLFSWCAFEV-----NAKYSVDVC 277
 QY 284 LFTDEWIOYDYLQSLGKYCYGAGNPLGPAQGVGFVNEILARLTHSPVODHTSTNHTLD 343
 Db 278 IPTKDELVHSYQDLHTYYHGGPGYDIKSVGNLNFASVKLLKQSEIQDQ----- 329
 QY 344 SNAPTFPLNATLYADFSDHNTVSWIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPF 403
 Db 330 -----KWLSFTHTDILNFLTLAGIIDDKNLNTAEYVPMGMT--FHRSWYVPQ 377
 QY 404 AARAVEVMQCEAEKPELVRLVNDVRVPLHGGVDKLGCRKDDF-----VEGLSFA 456
 Db 378 GARVYTERFQC--SNDTYRVYVINDAVVPIETCTSGPGFSCSEINDFYDYAEKRVAGTDFL 435
 QY 457 R 457
 Db 436 K 436

RESULT 5
 PPAB_YEAST

ID AC PPAB_YEAST STANDARD; PRT; 467 AA.
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE ACID PHOSPHATASE PHO11 PRECURSOR (EC 3.1.3.2) (P56).
 GN PHO11 OR YHR071W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=95249563; PubMed=7731988;
 RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
 RA Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
 RA Storms R.K.
 RT "The nucleotide sequence of chromosome I from Saccharomyces
 RT cerevisiae."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen J.Y., Gong Y.I., Ao S.Z.
 RT "The primary structure of acid phosphatase gene PHO11 in S.
 RT cerevisiae and comparison with other gene families."
 RL Acta Biochim. Biophys. Sin. 21:437-444(1989).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
 CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; L28920; AAC09508.1;
 DR EMBL; U19789; AAA73479.1;
 DR HSSP; P34752; LIHP.
 DR SGD; S0000094; PHO11.
 DR InterPro: IPR000560;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 461 461
 FT CONFLICT 17 17
 FT CONFLICT 82 82
 FT CONFLICT 150 150
 FT CONFLICT 354 354
 FT CONFLICT 423 423
 FT SEQUENCE 467 AA; 52757 MW; AECDC1C046B326C3 CRC64;

Query Match 16.2% Score 399; DB 1; Length 467;
 Best Local Similarity 26.1% Pred. No. 1e-24;
 Matches 121; Conservative 63; Mismatches 203; Indels 76; Gaps 14;
 QY 24 PRGNHSCDVTVDGQYQCFPEISHLWQYSPFSLADESAISPDVPGKCRVTFVQVLSRHG 83
 DB 21 PLGLSIDIKIGTQTEIFPFL---GGSGPVYSPGDYGISRDLPESCEKQVOMVGRHG 76
 QY 84 ARYPTSSKSKYSALIEAIQKNATAFKGKAFLEK-----TNYNTL-----GADDL 128
 DB 77 ERYPTVSKAKSIMTWTYKLSNYTGQFSGALSLFNDDEYEFFIRDTKNELETTLANSVNV 136
 QY 129 TPF-GEQVMVNSGIFKRYRKALARKLVPF--VRASGSDRVITASEKFTIEGFSQAKLADPG 186
 DB 137 NPYTGEMNAKRHARDFLAQYGVYVENQTSFAVTSNSNRCHDTAQYFIDG-----LGDK- 190
 QY 187 ANPHQASPVNLVII-----PEGAGYNNTLDHGLCTAFEESELGDDVDEANFTAVFAPPIRA 241
 DB 191 -----FNISLOTISEAESAGANTLSAHSAPWDD--DVNDDILKKYDTKYLGLGIAK 240
 QY 242 RLEAHLPGVNLTDDEVVNLMDCPEDIVARTSDATOLSPEDLFTHDWEIOYDYLQSLGK 301
 DB 241 RLNKENGLNLITSSDANTFCAWCAEINAR-----GYSDICNIFTKDELVRFSYGQDLET 295
 QY 302 YVYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADESH 361
 DB 296 YVYTGPGYDVRVSGANLFLNASVKLLKESEVDQ-----KWLISFTH 337
 QY 362 DNTMVSIFPALLGYNCTKPLTSTSVESIEETDGYASWVTPFAARAYVEMOCEAEKPL 421
 DB 338 DTDILNLYLTIGLIIDKKNLTAEHPFMENT--FHSWYVPOGARVYTEKFOC--SNDTY 393
 QY 422 VRVLVNDRVVPLHGGGVDRKGRKDDF-----VEGLSFAR 457
 DB 394 VRVINDAVVPIETCTGPGFSCINDEVDYAEKRVAGTDFLK 436
 RESULT 6
 PPAC_YEAST
 ID PPAC_YEAST STANDARD; PRT; 467 AA.
 AC P38693;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE ACID PHOSPHATASE PHO12 PRECURSOR (EC 3.1.3.2).
 GN PHO12 OR YHR215W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII."
 RL Science 265:2077-2082(1994).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
 CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

QY 109 FKGYAFLEK-TYNTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFUNDIYEFPIRDODDLEMETTFANSNDVNLNPTGEMDAKRHAREFLAQGYWFE 161
 QY 153 KIVPF-VRASGSDRVASAEKFIQFQSAKLADPGANPHQASPVNLIIPGAGYNNITLD 211
 Db 162 NOTSFPIFAASSERVHTAQYFIDGL-----GDQFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFESESELGDDVEANTAVFAPPARLEAHLPGVNLTDVVDVMDMCPFDIVA 270
 Db 211 AGNACPDWE-DANDDLDKDYDTYLDLIAKRLNKENKGLNLTAKDANTLFAWCAYELNA 269
 QY 271 RTSDATOLSPFCDLFTHDEWIOYDLOSGLKRYGYGAGNPLGPAQGVGVNELLARLTHS 330
 Db 270 R-----GYSVDCDIFTEDELVRYSYGODLVSYFGDGYDMIRSVGANLNFATLKLLKQS 324
 QY 331 PVQDHTSTNHLDSNPATFPLNATLYADFSDHNTMVSIFFGALGLYNGTKPLSTSVESIE 390
 Db 325 ETQD-----LKVLSFTHDIDLNLTTAGIIDDKNNLTAEYVVFMG 366
 QY 391 ETQYAASTVTPFAARAYVEMMOCEAEKEPLVRVLVNDRVVPLHGGCVGDKGRCKRDDF- 449
 Db 367 NT--FHKSWSVPOGARVYTEKFC--SNDTYVRVINDAVVPIETCSTGPGFSCINDFY 422
 QY 450 -----VEGLSFAR 457
 Db 423 DYAEKRVAGTDFLK 436

RESULT 8

PPAD_YEAST STANDARD; PRT: 468 AA.
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROBABLE ACID PHOSPHATASE YDL024C PRECURSOR (EC 3.1.3.2).
 GN YDL024C OR D2815.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Vissers S., Urrestarazu L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL; Z48432; CAAB8335.1; -
 CC DR EMBL; Z74072; CAAB9583.1; -
 CC DR SGD; S0002182; YDL024C.
 CC DR InterPro; IPR000560; -
 CC DR Pfam; PF00328; acid_phosphat_1; 1.
 CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC KW Hydrolase; Glycoprotein; Signal; Multigene family.
 CC SIGNAL 1 20 BY SIMILARITY
 CC CHAIN 21 468 PROBABLE ACID PHOSPHATASE YDL024C.
 CC FT ACT_SITE 76 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 CC FT ACT_SITE 338 338 PROTON DONOR (BY SIMILARITY).
 CC FT CARBOHYD N-LINKED (GLCNAC...) (POTENTIAL).
 CC

FT CARBOHYD 163 163 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDF162C CRC64;
 Query Match 15.08; Score 370; DB 1; Length 468;
 Best Local Similarity 25.18; Pred. No. 2.2e-22;
 Matches 116; Conservative 62; Mismatches 174; Indels 110; Gaps 14;
 QY 49 GOKSPFFSLADESAISPDVPGKCRVTEFVQLSRHGRARYPTSSKSKYSALTEAIQKNATA 108
 Db 43 GGSAPYFSPFANYGIPDTIPEGCRLTQVMIGRHGRYPTRSEAKDIFEVYKISNTYTK 102
 QY 109 FKGYAFLEK-TYNTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFUNDIYEFPIRDODDLEMETTFANSNDVNLNPTGEMDAKRHAREFLAQGYWFE 161
 QY 109 FKGYAFLEK-TYNTLGADD-----LTPF-GEQMVNSGKIFRYRYKALAR 152
 Db 102 FNGSLFUNDIYEFPIRDODDLEMETTFANSNDVNLNPTGEMDAKRHAREFLAQGYWFE 161
 QY 153 KIVPF-VRASGSDRVASAEKFIQFQSAKLADPGANPHQASPVNLIIPGAGYNNITLD 211
 Db 162 NOTSFPIFAASSERVHTAQYFIDGL-----GDQFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HG-LCTAFESESELGDDVEANTAVFAPPARLEAHLPGVNLTDVVDVMDMCPFDIVA 270
 Db 211 AGNACPDWE-DANDDLDKDYDTYLDLIAKRLNKENKGLNLTAKDANTLFAWCAYELNA 269
 QY 271 RTSDATOLSPFCDLFTHDEWIOYDLOSGLKRYGYGAGNPLGPAQGVGVNELLARLTHS 330
 Db 270 R-----GYSVDCDIFTEDELVRYSYGODLVSYFGDGYDMIRSVGANLNFATLKLLKQS 324
 QY 331 PVQDHTSTNHLDSNPATFPLNATLYADFSDHNTMVSIFFGALGLYNGTKPLSTSVESIE 390
 Db 325 ETQD-----LKVLSFTHDIDLNLTTAGIIDDKNNLTAEYVVFMG 366
 QY 391 ETQYAASTVTPFAARAYVEMMOCEAEKEPLVRVLVNDRVVPLHGGCVGDKGRCKRDDF- 449
 Db 367 NT--FHKSWSVPOGARVYTEKFC--SNDTYVRVINDAVVPIETCSTGPGFSCINDFY 422
 QY 450 -----VEGLSFAR 457
 Db 423 DYAEKRVAGTDFLK 436

RESULT 9

PHYB_ASPNG STANDARD; PRT: 479 AA.
 AC P34754;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE B).
 GN PHYB.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.
 RX MEDLINE=93371452; PubMed=7916610;
 RA Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,
 RA Ullah A.H.J.;
 RT "Identification and cloning of a second phytase gene (phyB) from
 FT Aspergillus niger (ficcum).";
 FT Biochem. Biophys. Res. Commun. 195:53-57(1993).
 RL

CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L20567; AAA02934.1; -
 CC InterPro: IPR000560; -
 CC Pfam: PF00328; acid_phosphat; 1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal.
 CC SIGNAL 1 19
 CC CHAIN 20 479 3-PHYTASE B.
 CC ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE
 CC (BY SIMILARITY).
 CC ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 CC ACT_SITE 337 337 PROTON DONOR (BY SIMILARITY).
 CC CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 479 AA; 52611 MW; 395D4DA2B50DFC4 CRC64;
 SQ
 Query Match 14.3%; Score 352.5; DB 1; Length 479;
 Best local Similarity 27.9%; Pred. No. 5.9e-21;
 Matches 117; Conservative 54; Mismatches 158; Indels 91; Gaps 16;
 QY 49 GOYSPFFSLADESAISPDVPRGCRVTEVQVLSRHGARYPTSSKSKY-SALIEAOKNAT 107
 Db 53 GPYSERSVY---GIARDPTGCEVDQVIMVKRIGERYSPSAGKSTEEALAKVYSINT 108
 QY 108 AFKGYAPLKYNTL-----GADDLT-PF-GBOQMVNSGIKRYRYKAL--ARKIVPF 157
 Db 109 EYKGLAFLNDWTYVNECYNAETTSPPVAGLLDAYNHGNDYKARYGHLWNGEIVVVF 168
 QY 158 VRAGSDRVIASAEKFIQFOSAKLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTA 217
 Db 169 F-SSGGRVETARKFGEF-----DVEANFTAVFAPPR--ARLEAHLPGVNTLDDVNLMD 262
 QY 218 FEESELG-----DVEANFTAVFAPPR--ARLEAHLPGVNTLDDVNLMD 262
 Db 202 SESEVMGADSLTPTCDTNDQTTCDNLTYQLPQFKVAARLNSQNGMNTASDVYVNLIV 261
 QY 263 MCPEDTARTSDATQSPFCDLFTHEWIOYDLOSLKYYGAGNPLGPAQGVGVNE 322
 Db 262 MASFEINAR-----PFSWNAFTQDEWVSFGYVEDLNYYICAGPKDNMAAVGAVYANA 316
 QY 323 LIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADESHDNTWYSIFALGLY--NGTKP 380
 Db 317 SLTLLNQG-----KEAGP-----LFFNFAHDTNITPILAAALGVLPNEIDL 358
 QY 381 LSTTSVESIEETDGYAASWTVPFAARAYVEMOCEA-----EKEPLRVLVNDRVPLHGC 436
 Db 359 L-----DRVAFGNPYIGNIVPMGGLHILIERLSQCATALSQKGYVRLVINEAVLPENDC 413
 RESULT 10
 PPAL_PICPA STANDARD; PRT; 468 AA.

AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ACID PHOSPHATASE PHOI PRECURSOR (EC 3.1.3.2).
 GN PHOI.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E., Gannon P.M., Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RL Gene 163:19-26(1995).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: BY PHOSPHATE STARVATION.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U28658; AAA85503.1; -
 CC HSSP; P34752; LIHP.
 CC InterPro: IPR000560; -
 CC Pfam: PF00328; acid_phosphat; 1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; Multigene family.
 KW SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 468 ACID PHOSPHATASE PHOI.
 FT ACT_SITE 84 84 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 345 345 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 52690 MW; AE555B8E27718C2C CRC64;
 Query Match 14.2%; Score 350; DB 1; Length 468;
 Best local Similarity 25.4%; Pred. No. 9.1e-21;
 Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;
 QY 6 VLIISIATL---FGSTSGTALGPRGNHSDTVDGYYQCFPEISHLWQYSPF--FSLADE 60
 Db 10 IILATLQSVFAVELQHVLYGVNDYRYPQRT-DDQYNILRHLEGL- GPYIGYNGWGIAAE 67
 QY 61 SAISPDVPRGCRVTEVQVLSRHGARYPTSSKSKYSALIE-AIOKNATAFKGYAFLKTY 119
 Db 68 SEI-----ESCTIDQAHLLMRHGERYPSTNVGKQLEALYQKLLDADVEPTGPLSFQDY 122
 QY 120 NYTLG-----ADDLTPFGEQMVNSGIKF---YRRYKAL-----ARKIVPFVRASGSDR 165
 Db 123 DYFVSDAAAYEQETTKGFFSGLNTAFDFTTLRERYDHLINTSECKKL--SWWAGSOER 180
 QY 166 VIASAEKFIQFOSAKLADPGANPHQASPVINVIPE---GAGYNNTLDHGLCT-----A 217
 Db 181 VDTAKYFAQCGKMSNYTD-----MVEVVALEEEKSGLNSLTARISCPNVTNSH 230
 QY 218 FESESELGDDV---EANFTAVFAPIRARLEAHLPGVNTLDDVNLMDMCPEDTARTSD 274
 Db 231 YKGDFFPDIAEREAD-----RLNTLSFGFNITADDIPTIALYCGFELNVGRG- 278

QY 275 ATQLSPFCDLTHDEWIOYDIYLSQKGYGAGNPLGPAQGVGVNBLIARLTHSPVQD 334
 DB 279 ----SSFCVDLSREALITYALRDUGVYVNGNPLGKTIYGVYAN 321
 QY 335 HTSTNHTLDSNPATPLNATLYADSHONTMVSIFPAGLYNGTK-PLSTTSVRSIEBTD 393
 DB 322 --ATQLLENTAD-PRDPLYSFSDHDTLLQVFTSLGLENVTDLPL-----DOIQFOT 373
 QY 394 GYAASWTVPFAARAVAMQC--EAEKPLVRLVNDKRGVVDKLGRCRKRDDFEV 451
 DB 374 SPKSTFVPMGARLLTERLLCTVEGEKYYVYRTILNDVAFPLSDCSCGPGFSCPLNDVYS 433
 QY 452 GLSFARSGGNWEE 464
 DB 434 RLEALNEDSFAE 446

RESULT 11
 PHYB ASPAW
 ID PHYB ASPAW STANDARD; PRT; 479 AA.
 AC P34755;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (PH 2.5 OPTIMUM ACID PHOSPHATASE).
 GN PHYB OR APH.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 ON NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ALK0243;
 RX MEDLINE=94040796; PubMed=8224894;
 RA Fiddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
 RA Miettinen-Ginonen A., Nevalainen H., Rambosek J.A.;
 RT "The cloning and sequencing of the genes encoding phytase (phy) and
 RT PH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
 RT awamori.";
 RL Gene 133:55-62(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=99264417; PubMed=10329192;
 RA Kostrewa D., Wyss M., D'Arcy A., van Loon A.P.;
 RT "Crystal structure of Aspergillus niger pH 2.5 acid phosphatase at
 RT 2.4-A resolution.";
 RL J. Mol. Biol. 288:965-974(1999).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L02420; AAA16897.1;
 CC PIR: JN0890; JN0890.
 CC DB: IQFX; 19-APR-00.
 CC InterPro: IPR000560;
 CC Pfam: PF00328; acid_phosphat; 1.
 CC PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; 3d-structure.
 CC SIGNAL 1 19 BY SIMILARITY.

FT CHAIN 20 479 3-PHYTASE B.
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR.
 FT ACT_SITE 337 337 PROTON DONOR.
 FT DISULFID 71 387
 FT DISULFID 128 472
 FT DISULFID 216 441
 FT DISULFID 225 298
 FT DISULFID 413 421
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 479 AA; 52678 MW; 4F8E0F3778CC3B08 CRC64;

Query Match 14.0%; Score 346.5; DB 1; Length 479;
 Best Local Similarity 27.4%; Pred. No. 1.8e-20;
 Matches 115; Conservative 54; Mismatches 160; Indels 91; Gaps 15;

QY 49 GOYSPFFSLADESAISPDVPKGRVTEQVLSRHRGARYPTSSKSKY-SALIEATQKNAT 107
 DB 53 GYSPERSVY---GIARDPPTSCVDQVIMVKKRGERIPSPSAGKDIEELAKVYSINT 108
 QY 108 AFGKYAEFLKTYNTL-----GADLT-PF-GEQMVNSGKIFKRYKAL--ARKIVPF 157
 DB 109 EYKGLAFLNDWYVYPNECYNAETTSQPYAGLLDAYNHGNDYKARYGHLWNGETVVPF 168
 QY 158 VRAGSDRVIASAEKFIQGFQSAKLADPGANPHQASPVINIIPEGAGYNNVTLDHGLCTA 217
 DB 169 F-SSQYGRVETARKEGFGF-----FGYNYSTNAALNII 201
 QY 218 FESEELG-----DDVEANFTAVFAPPRI--ARLEAHLPGVNLTDVVDVNLMD 262
 DB 202 SESEVMGADSLTPTCDTNDQTTCDNLTYQLPQFKVAARLNSQNGMNLTSADVYMLAV 261
 QY 263 MCPEDTVAATSDATQLSPFCDLTHDEWIOYDIYLSQKGYGAGNPLGPAQGVGVNE 322
 DB 262 MASFELNAR-----PFSNWINAFTQDEWVSFGYVEDLNYYCYAGPGDKNMAAVGAVYANA 316
 QY 323 LIARLTSPVDHTSTNHTLDSNPATPLNATLYADSHONTMVSIFPAGLY--NCTKP 380
 DB 317 SLATLLNQGPKE-----AGSLFFNFPAHDTNITPILAAUGVLIPNEDLP 358
 QY 381 LSTTSVESIEETDGYAASWTVPFAARAVYEMMOCEA-----EKEPLVRLVNDKRGVVD 436
 DB 359 L-----DRVAFGNPYSIGNIVPMGGLHTIERLSQATLSDEGTYVRLVNEAVLPFND 413

RESULT 12
 PPAS_KLULA
 ID PPAS_KLULA STANDARD; PRT; 469 AA.
 AC P52289;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHOS.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RA Ferminan E.;
 RL Thesis (1995). University of Salamanca, Spain.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

Db 82 DAVGTANFQORLLGVSVPIDYVSGNPLSFVPTWTFVIEAANADALSSGRVFLDMGRQ 141
 QY 143 FYRRYKALARKIVPVRASGSDRVITASAEKTEGFSQAKLADPGANPHOASPVINVIPIE 202
 Db 142 FYERHFNASTYNIYTAQORVVSALWTGYGMF-----GEDVFNFTNYILVSENA 194
 QY 203 GAGYNTLDHGLCTAFESSELGDDVEANFTAVFAPPIRARLEAHLPGVNLDEDVNLMD 262
 Db 195 TAGSNLSYNACASDADDTTTPALEARNVYMEPIRQRLNPNFYSNLYNLNDLLNLIG 254
 QY 263 MCPEDTVAARTDAQLSPFCDLFTHDEWIQDYLOSGLKYYGYGAGNPLGPAQGVGVNE 322
 Db 255 ICSVEIALQ-----DYSEFKCLFNSVDFLNFPEYEGDLFSFSGMNSVKWGSIFGGAYANS 309
 QY 323 LIARLTHSPVQDHTSTNHTLDSNPATEPLNATLADSHDNTWYSIFFALGLYNGTKPLS 382
 Db 310 LANSLS-RSVENNTQ-----QVFFATHDANIIPVETALGFFDNTPEN 351
 QY 383 TTSVESIEETDGYAASVTPFAARAYVEMMOCEAEKEPLRVLRVLVNDRVVPLHGCG 437
 Db 352 PLPTSQVHSHMSKASEFVPFAGNLITELFCEDSKY-YVRHLVNEEVPLSDCG 405

RESULT 14

PPA2_SCHPO STANDARD; PRT; 463 AA.
 AC Q01682;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE THIAMINE-REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO4 OR SPEC428.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
 RX MEDLINE=91064763; PubMed=2249257;
 RA Yang J., Schweingruber M.E.;
 RT "The structural gene coding for thiamin-repressible acid phosphatase
 in Schizosaccharomyces pombe";
 RL Curr. Genet. 18:269-272(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- INDUCTION: REpressed BY THIAMINE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X56939; CAA0258.1; --
 CC EMBL; AL034382; CAA2278.1; --
 CC PIR; S14119; S14119;
 CC InterPro; IPR000560;
 CC Pfam; PF00328; acid.phosphat.1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Cell wall; Signal.

FT SIGNAL 1 18
 FT CHAIN 19 463
 FT ACT_SITE 69 69
 FT ACT_SITE 340 340
 FT CARBOHYD 98 98
 FT CARBOHYD 104 104
 FT CARBOHYD 186 186
 FT CARBOHYD 221 221
 FT CARBOHYD 251 251
 FT CARBOHYD 328 328
 FT CARBOHYD 433 433
 FT CARBOHYD 439 439
 FT CARBOHYD 458 458
 SQ SEQUENCE 463 AA; 52118 MW; F48EAF8BB6B234A CRC64;

Query Match 12.28; Score 301; DB 1; Length 463;
 Best Local Similarity 25.4%; Pred. No. 8.1e-17;
 Matches 105; Conservative 66; Mismatches 185; Indels 58; Gaps 14;

QY 51 YSPFFSLADESAISPDVPKGRVTFVQVLSRHGARYPTSS-----KSKKYSALIE----- 100
 Db 43 HEYPF-----NGPTTFPESCAIKQVHLLQHGSRNPTGDDTATDVSSAQIDIFQNKLL 97
 QY 101 --AIOKNATAFKGYAFKTYNYTL---GADDLTPFGQQMVNSGIRYRYKALARKIV 155
 Db 98 NGSLPVNFSPENPLFVKHWTVPVKAENADQLSSSGRIELFDLGRQVFRYYELFDTDV 157
 QY 156 PFVRASGSDRVITASAEKIFGFSQAKLADPGANPHOASPVINVIPIE--GAGYNTLDHG 213
 Db 158 YDINTAAQERVDASAEEFSGM-----FGDDMQNTNF-----IVLPEDDSAGANSLAMY 208
 QY 214 LCTAFEESELGDD-VEANFTA---VFAPPIRARLEAHL-PCVNLTDSDVNLMDMCPEDT 268
 Db 209 SCPVYEDNDNIDENTEAHTSWRNVFLKPIANLNKYFDSGYNLTVSDVRSLYICVYEI 268
 QY 269 VARTSDATQLSPFCDLFTHDEWIQDYLOSGLKYYGYGAGNPLGPAQGVGVNELIARLT 328
 Db 269 ALRDN-----SDFCSLFTPEFLNFYDSDLDYAYWGGPASEWASTLGGAYVNNLANNL- 322
 QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTWYSIFFALGLYNGTKPLSTSVES 388
 Db 323 -----RKGVNNSDRK-----VFLAFTHDSQIIPVEAALGFDPDITPHEPLPTDK 367
 QY 389 IBETDGYAASVTPFAARAYVEMMOCEAEKEPLRVLRVLVNDRVVPLHGCGVDKLG 442
 Db 368 NIFTYSLKTSFVPFAGNLITELFLC-SDNKYYVVRHLVNVQVVPVPLTDCGVGPSG 420

RESULT 15

PPAX_CAEEL STANDARD; PRT; 755 AA.
 AC Q09549; Q17843;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE ACID PHOSPHATASE F26C11.1 (EC 3.1.3.2).
 GN F26C11.1
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Matthews P., Lloyd C.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Sat Oct 27 15:25:41 2001

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

DR EMBL; Z47072; CAA87370.1; -
DR EMBL; Z54342; CAA87370.1; JOINED.
DR EMBL; Z54342; CAA91156.1; -
DR EMBL; Z47072; CAA91156.1; JOINED.
DR WormPep; F26C11.1; CE05732.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein; Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
FT SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.4%; Score 134; DB 1; Length 755;
Best Local Similarity 21.9%; Pred. No. 0.0049;
Matches 91; Conservative 56; Mismatches 123; Indels 146; Gaps 23;
QY 57 LADSAISPDVPKGVKRVTFVQVLSRHGAR-----YPTSSSKSKYSALIEATQKNATAFK 110
Db 332 LAIEKFMSEV-DNMKLEFVQIWRHGRDRALEGLFPISEK----- 371
QY 111 GKYAFLLKTYNTLGA---DDLTPFGQOMVNSGKPYRY-----KALARKIVPEV 158
Db 372 -----NWTFGGGGLGELTPMGSEMNNLGTIFRRRYVEDQOFLSHRYAAKEI--YI 420
QY 159 RASGSDRVIAAEKFTGFSQAKLADPGANPHQASPVINVIIEGAGYNNITLD----- 211
Db 421 RSTNLNLTIIISAMSLYG-----MFPFGA-----WNI---QGVDPNDVDWQOGFTF 464
QY 212 -----HGL-----CTAFEE-----SELGDDVEANFTAVFAPPFIRARLEAHLPGVN 251
Db 465 IPHVVDGIDQCAVALCNCRRFOELQEKWAEI-DEYKNATVAMIA--LNRRVAAF---YN 518
QY 252 LTDEDVNLMDMCPFFDVARTSDA--TQLSPFCD-LFTHDEWIOYDYLQSLGKYG---- 304
Db 519 VTQDP-----EKFNRYTDARKCORWFNDTWYQQLPWYNEDLYNEAQRYAPEKR 568
QY 305 ---YGAGNP-----LGPAGGVGVNELLARLTHSPVQDHTSTNHTLD--SNPA 347
Db 569 FTEGFGNPKPSIVDGDIDPQEVSTLQGGPLLNEIFERGKIRCVADAENCSDIYLKPL 628
QY 348 TFPLNATLYADFSDHNTWVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPF 403
Db 629 KF-----YAYSSHDLQVYALLVTGI-----TDVKTVDGWPDTSLSLTIEY 670

Search completed: October 26, 2001, 15:18:02
Job time: 448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: October 26, 2001, 15:18:09 ; Search time 32.57 Seconds
(without alignments)
295.232 Million cell updates/sec
Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGVPVLLSIATLFGSTSGT.....DFVEGLSFARSGGNWEECPA 467
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 197339 seqs, 20590346 residues
Total number of hits satisfying chosen parameters: 197339
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Issued_Patents_AA:**
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:**
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:**
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:**
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:**
5: /cgn2_6/ptodata/2/iaa/6C_COMB.pep:**
6: /cgn2_6/ptodata/2/iaa/6D_COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2185	88.5	441	4	US-09-121-425-1
2	2173	88.0	467	4	US-09-121-425-2
3	1866	75.5	467	1	US-07-923-724-8
4	1866	75.5	467	2	US-08-609-426A-8
5	1866	75.5	467	2	US-08-374-652C-2
6	1862	75.4	467	1	US-08-151-574-32
7	1862	75.4	467	1	US-08-146-424-20
8	1862	75.4	467	1	US-08-693-709-2
9	1862	75.4	467	2	US-08-419-448-32
10	1862	75.4	467	2	US-08-819-825-3
11	1862	75.4	467	4	US-09-163-642-3
12	1855	75.1	467	4	US-09-155-855-3
13	1813	73.4	443	4	US-09-155-855-1
14	1808	73.2	443	4	US-09-155-855-2
15	1395	56.5	475	2	US-08-819-825-2
16	1395	56.5	475	2	US-08-819-825-2
17	846.5	34.3	443	3	US-08-993-359-30
18	833.5	33.7	453	3	US-08-993-359-22
19	831.5	33.7	439	3	US-08-993-359-24
20	831.5	33.7	439	3	US-09-221-654-2
21	831.5	33.7	439	3	US-08-989-358A-2
22	822	33.3	442	3	US-08-993-359-28
23	811	32.8	442	3	US-08-993-359-26
24	359	14.5	468	1	US-07-627-539G-2
25	354.5	14.4	446	1	US-07-627-539G-7
26	346.5	14.0	479	1	US-07-923-724-2
27	346.5	14.0	479	2	US-08-609-426A-2

28	346.5	14.0	479	2	US-08-374-652C-4
29	185.5	7.5	92	3	US-08-993-359-32
30	179	7.2	449	3	US-08-680-506-7
31	121.5	4.9	318	3	US-08-680-506-3
32	121	4.9	113	1	US-08-241-853-8
33	121	4.9	113	1	US-08-850-917-8
34	115	4.7	24	2	US-08-374-652C-32
35	110	4.5	113	1	US-08-241-853-10
36	110	4.5	113	2	US-08-850-917-10
37	98	4.0	386	1	US-08-758-213-1
38	98	4.0	386	2	US-08-692-787-48
39	98	4.0	386	4	US-09-097-199-48
40	98	4.0	515	2	US-09-146-283-2
41	98	4.0	515	3	US-08-579-823A-2
42	98	4.0	515	4	US-09-344-195-2
43	97	3.9	35	1	US-08-151-574-2
44	97	3.9	35	2	US-08-419-448-2
45	93.5	3.8	675	4	US-09-171-878-1

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-1

Query Match	88.5%	Score 2185;	DB 4;	Length 441;
Best Local Similarity	92.7%	Pred. No. 3.6e-222;		
Matches	409;	Conservative	13;	Mismatches 19;
		Indels	0;	Gaps 0;
QY	27	NSHSCDVTGQYQCFPEISHLWGQYSPFFSLADESALSPDVPKGRVTFVOVLSRHGARY	86	
Db	1	NSHSCDVTGQYQCFPEISHLWGQYSPFFSLADESALSPDVPKGRVTFVOVLSRHGARY	60	
QY	87	PTSSKSKYSALIEAIOKNATAFKGYAEIKTYNYLGLADLTTPFGEQMVNSGIFKFFRR	146	
Db	61	PTSSKSKYSALIEAIOKNATAFKGYAEIKTYNYLGLADLTTPFGEQMVNSGIFKFFRR	120	
QY	147	YKALARKIVPFVPCASSDRVIAEAEKIEGFSQSAKLADPGANPHOASPVINVIIPEGAGY	206	
Db	121	YKALARKIVPFVPCASSDRVIAEAEKIEGFSQSAKLADPGANPHOASPVINVIIPEGAGY	180	
QY	207	NNTLDHGLCTAFESSELGDDVEANFTAFPIRLEAHLPGVNLTDDEVVNLMDCPF	266	
Db	181	NNTLDHGLCTAFESSELGDDVEANFTAFPIRLEAHLPGVNLTDDEVVNLMDCPF	240	
QY	267	DTVARTSDATQLSPCDLFTHDEWTOYDYLQSLGKYYGYGAGNPLGPAQGVGFNELLAR	326	
Db	241	DTVARTSDATQLSPCDLFTHDEWTOYDYLQSLGKYYGYGAGNPLGPAQGVGFNELLAR	300	
QY	327	LPHSPVDHTSTNTFLDNPATFPLNATLYADFSHDNTMVSIFPFGALGYNKTKPLSTTSV	386	

Sat Oct 27 15:25:39 2001

Db 301 LTRSPVQDHTSTNHTLDSNPATFFPLNATLADFSHDSNMISIFFALGLYNGTAPLSTTSV 360
QY 387 ESIEETDGYAASWTVPFAARAYVEMMOCEAEKPELVRLVNDVRVPLHGGVCKLGRCKR 446
Db 361 ESIEETDGYASWTVPFCARAYVEMMOCEAEKPELVRLVNDVRVPLHGGCAVCKLGRCKR 420
QY 447 DDFVEGLSFARSGGNWECFA 467
Db 421 DDFVEGLSFARSGGNWECFA 441
RESULT 2
US-09-121-425-2
; Sequence 2, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2
Query Match 88.0%; Score 2173; DB 4; Length 467;
Best Local Similarity 85.2%; Pred. No. 7.3e-221; Indels 40; Gaps 2;
Matches 415; Conservative 13; Mismatches 19;
QY 1 MGVEVLLSTATLFGSTGTALPGRGNHSCDVTGQYQCFPEISHLMGQYSPFFSLADE 60
Db 1 MGVEVLLSTATLFGSTGTALPGRGNHSCDVTGQYQCFPEISHLMGQYSPFFSLEDE 60
QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAPLKTYN 120
Db 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAPLKTYN 120
QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
Db 101 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIAAEKFIQFQSA 160
QY 181 KLADPGANPHQASPVIN-----VITPEGAGYNNTLDHGLCTAFEE 220
Db 161 KLADPGSOPHQASPVIDILIEAIQKNATAFKGYAPLKTYN-----TYN 100
QY 221 SELGDDVEANFTAFAPPIRLEAHLPGVNLTDDEDVNLMDMCPPTVARTSDATQLSP 280
Db 221 SELGDDVEANFTAFAPPIRLEAHLPGVNLTDDEDVNLMDMCPPTVARTSDATQLSP 280
QY 281 FCDLFTHDWIOYDYLQSLGKYGYGAGNPLGPAQGVGVNELLARTHSPVDHTSTNH 340
Db 281 FCALFTHDWIRQYDYLQSLGKYGYGAGNPLGPAQGVGVFANELARTLRSPVDHTSTNH 340
QY 341 TLDSPATPLNATLYADFSHNTWVSTIFFALGLYNGTKPLSTTSVESIETDGYASWT 400
Db 341 TLDSPATPLNATLYADFSHNTWVSTIFFALGLYNGTKPLSTTSVESIETDGYASWT 400
QY 401 VPFAARAYVEMMOCEAEKPELVRLVNDVRVPLHGGVCKLGRCKRDFVEGLSFARSGG 460
Db 401 VPFAARAYVEMMOCEAEKPELVRLVNDVRVPLHGGCAVCKLGRCKRDFVEGLSFARSGG 460
QY 461 NWEECFA 467
Db 461 NWEECFA 467

Db 461 NWEECFA 467
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerström, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07923,724
; APPLICATION NUMBER: US/07923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8
Query Match 75.5%; Score 1866; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 1.9e-188; Indels 0; Gaps 0;
Matches 348; Conservative
QY 1 MGVEVLLSTATLFGSTGTALPGRGNHSCDVTGQYQCFPEISHLMGQYSPFFSLADE 60
Db 1 MGVEVLLSTATLFGSTGTALPGRGNHSCDVTGQYQCFPEISHLMGQYSPFFSLANE 60
QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAPLKTYN 120
Db 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAPLKTYN 120

QY 121 YTLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKTFIEGFQSA 180
 Db 121 YSLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKTFIEGFQSA 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 QY 241 ARLEAHLPGVNLDEDEDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIDYDYLQSLG 300
 Db 241 ORLENDLSGVTLTDETVYLMDCMCFDFTSTVDTKLSFPCDLFTHDEWIDYDYLQSLG 300
 QY 301 KYGYGAGNPLGACQGVFNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 Db 301 KYGYGAGNPLGACQGVFNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 QY 361 HNTMWSIFFALGLYNGTKPLSTTSVESTTEEDGYAAWTVPFARAYVEMMOCEAKEP 420
 Db 361 HNTMWSIFFALGLYNGTKPLSTTSVESTTEEDGYAAWTVPFARAYVEMMOCEAKEP 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467

RESULT 4

US-08-609-426A-8
 ; Sequence 8, Application US/08609426A
 ; Patent No. 5830733
 ; GENERAL INFORMATION:
 ; APPLICANT: Nevalainen, Helena K.M.
 ; APPLICANT: Paloheimo, Marja T.
 ; APPLICANT: Miettinen-Oinonen, Arja S.K.
 ; APPLICANT: Torkkeli, Tuula K.
 ; APPLICANT: Cantrell, Michael
 ; APPLICANT: Piddington, Christopher S.
 ; APPLICANT: Rambosek, John A.
 ; APPLICANT: Turunen, Marja K.
 ; APPLICANT: Fagerstr m, Richard B.
 ; APPLICANT: Houston, Christine S.
 ; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 01-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/923,724
 ; FILING DATE: 31-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/496,155
 ; FILING DATE: 19-MAR-1990
 ; APPLICATION NUMBER:
 ; FILING DATE: 29-APR-1987
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: UK 8610600
 ; FILING DATE: 30-APR-1986
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Reed, Grant E.

REGISTRATION NUMBER: P-41,264
 REFERENCE/DOCKET NUMBER: 1050.0080001
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-609-426A-8

Query Match 75.5%; Score 1866; DB 2; Length 467;
 Best Local Similarity 74.5%; Pred. No. 1.9e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGQYSPFFSLADE 60
 Db 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGQYQCFPEISHLWGQYSPFFSLADE 60
 QY 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSKKYSALIEATOKNATAFKGYAFKLTYN 120
 Db 61 SAISPDVPGCRVTFVQVLSRHGARYPTSSKKYSALIEATOKNATAFKGYAFKLTYN 120
 QY 121 YTLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKTFIEGFQSA 180
 Db 121 YSLGADLLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVITASAEKTFIEGFQSA 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPGANPHQASPVINVIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 QY 241 ARLEAHLPGVNLDEDEDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIDYDYLQSLG 300
 Db 241 ORLENDLSGVTLTDETVYLMDCMCFDFTSTVDTKLSFPCDLFTHDEWIDYDYLQSLG 300
 QY 301 KYGYGAGNPLGACQGVFNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 Db 301 KYGYGAGNPLGACQGVFNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 QY 361 HNTMWSIFFALGLYNGTKPLSTTSVESTTEEDGYAAWTVPFARAYVEMMOCEAKEP 420
 Db 361 HNTMWSIFFALGLYNGTKPLSTTSVESTTEEDGYAAWTVPFARAYVEMMOCEAKEP 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGVDKLGRCRDRDVEGLSFARSGGNWEECEFA 467

RESULT 5

US-08-374-652C-2
 ; Sequence 2, Application US/08374652C
 ; Patent No. 5834286
 ; GENERAL INFORMATION:
 ; APPLICANT: NEVALAINEN, HELENA K.M.
 ; APPLICANT: PALOHEIMO, MARJA T.
 ; APPLICANT: FAGERSTROM, RICHARD B.
 ; APPLICANT: MIETTINEN-OINONEN, ARJA S.
 ; APPLICANT: TURUNEN, MARJA K.
 ; APPLICANT: RAMBOSEK, JOHN A.
 ; APPLICANT: PIDDINGTON, CHRISTOPHER S.
 ; APPLICANT: HOUSTON, CHRISTINE S.
 ; APPLICANT: CANTRELL, MICHAEL A.
 ; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
 ; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
 ; CITY: WASHINGTON

us-09-488-265-26.ra1

Sat Oct 27 15:25:39 2001

```

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: PCT/US93/07058
APPLICATION NUMBER: 27-JUL-1993
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-374-652C-2

```

Query Match 75.5%; Score 1866; DB 2; Length 467;
 Best Local Similarity 74.5%; Pred. No. 1.9e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

```

QY 1 MGFWVLLSIATLFGSTGATLGRGNHSCDVTVDGQYQCFPEISHLWQYSPFFSLADE 60
DB 1 MGVSALLPLVLLAGVTSGLAVPASRNQSCDVTVDGQYQCFSETSHLWQYAPFFSLANE 60
QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
DB 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
QY 121 YTLGADDLTPFGQOMVNSGKIFVRRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
DB 121 YSLGADDLTPFGQOMVNSGKIFVRRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
QY 181 KLADPGANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVARTSDATQSPFCDLTFHDEWTOYDYLSLG 300
DB 241 ARLEAHLPGVNLTDDEVNLMDCPFDVARTSDATQSPFCDLTFHDEWTOYDYLSLG 300
QY 301 KYGCGAGNPLGPTQGVYANELLARLTHSPVQDHTSNHLDNSNPATFPLNATLYADF 360
DB 301 KYGCGAGNPLGPTQGVYANELLARLTHSPVQDHTSNHLDNSNPATFPLNATLYADF 360
QY 361 HDNMTSIFPALGLYNGTKPLSTTSVESIEEDGYAASVTPFAARAYVEMMCQAEKEP 420
DB 361 HDNMTSIFPALGLYNGTKPLSTTSVESIEEDGYAASVTPFAARAYVEMMCQAEKEP 420
QY 421 LVRVLNDRVVPVHGGGVKDLGRCKRDEVEGLSFARSGNNEECPA 467
DB 421 LVRVLNDRVVPVHGGGVKDLGRCKRDEVEGLSFARSGNNEECPA 467

```

```

RESULT 6
US-08-151-574-32
; Sequence 32, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-151-574-32

```

Query Match 75.4%; Score 1862; DB 1; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

```

QY 1 MGFWVLLSIATLFGSTGATLGRGNHSCDVTVDGQYQCFPEISHLWQYSPFFSLADE 60
DB 1 MGVSALLPLVLLAGVTSGLAVPASRNQSCDVTVDGQYQCFSETSHLWQYAPFFSLANE 60
QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
DB 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
QY 121 YTLGADDLTPFGQOMVNSGKIFVRRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
DB 121 YSLGADDLTPFGQOMVNSGKIFVRRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
QY 181 KLADPGANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVARTSDATQSPFCDLTFHDEWTOYDYLSLG 300
DB 241 ARLEAHLPGVNLTDDEVNLMDCPFDVARTSDATQSPFCDLTFHDEWTOYDYLSLG 300

```

QY 301 KYGYGAGNPLGPGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 DB 301 KYGYGAGNPLGPGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMVSIFFFALGNGTKPLSTTSVESIBETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGNGTKPLSTTSVESIBETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 QY 421 LVRVLVNDVRVPLHGCGVDKLGRCRDRDVEGLSFARSGNWECEFA 467
 DB 421 LVRVLVNDVRVPLHGCGVDKLGRCRDRDVEGLSFARSGNWECEFA 467

RESULT 7
 US-08-146-424-20
 ; Sequence 20, Application US/08146424
 ; Patent No. 5593963
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/146,424
 FILING DATE: 02-NOV-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KENNEDY, BILL
 REGISTRATION NUMBER: 33,407
 REFERENCE/DOCKET NUMBER: 44615-20011.24
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-146-424-20

Query Match 75.4%; Score 1862; DB 1; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;
 QY 1 MGTVFVLLSIATLFGSTSGTALGRGNSHSCDVTVDGYQCFFPELSHWGQYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRQSCDVTVDGYQCFFPELSHWGQYSPFFSLADE 60
 QY 61 SAISPDVPGKCRVTFVQVLSHRGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTN 120
 DB 61 SVTSPVPGACRVTFQVLSHRGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTN 120
 QY 121 YTLGADLLTPFGEGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180

DB 121 YSLGADLLTPFGEGQMVNSGKIFRYRYESLNRNIVFFIRSSGSSRVIASGKKFIEGFQST 180
 QY 181 KLADPGANPHQASPVNVIIPGAGYNNNTLDHGLCTAFEESELGDDVEANFTAVFAPDIR 240
 DB 181 KLKDPRAQPGQSSPKIDVVISSESSNNNTLDPGTCTVFEDSELADTVANFTATFVPSIR 240
 QY 241 ARLEAHLPGVNLTDREVVNLMDMCPEDTAVATSDATQLSPFCDLFTHDEWIOYDYLOSIG 300
 DB 241 QRLNDLSGVTLTDEVTYLMDCSFDISTSTVDTKLSPFCDLFTHDEWINYDYLOSIG 300
 QY 301 KYGYGAGNPLGPGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 DB 301 KYGYGAGNPLGPGVGVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMVSIFFFALGNGTKPLSTTSVESIBETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISILFALGNGTKPLSTTSVESIBETDGYAASWTPFAARAYVEMMOCEAEKEP 420
 QY 421 LVRVLVNDVRVPLHGCGVDKLGRCRDRDVEGLSFARSGNWECEFA 467
 DB 421 LVRVLVNDVRVPLHGCGVDKLGRCRDRDVEGLSFARSGNWECEFA 467

RESULT 8
 US-08-693-709-2
 ; Sequence 2, Application US/08693709
 ; Patent No. 5770413
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/693,709
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146,424
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20011.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-813-5600
 ; TELEFAX: 415-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal

us-09-488-265-26.rai

Sat Oct 27 15:25:39 2001

```

FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1...23
; OTHER INFORMATION:
US-08-693-709-2

Query Match      75.4%; Score 1862; DB 1; Length 467;
Best Local Similarity 74.5%; Pred. No. 5.2e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCFPEITSHLWQYSPFFSLADE 60
Db 1 MGVSAYLLPLLYLLSGVTSLGAVPASRNQSSCDTVDQGYQCFSEITSHLWQYAPFFSLANE 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFRASGSDRVIASAEKFEIGFQSA 180
Db 121 YSLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFRASGSDRVIASAEKFEIGFQSA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNILDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVINVIPEGAGYNNILDHGLCTAFEESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVDVNLMDMCPEDTIVARTSDATQSPFCDLFTHDWMIOYDLSLG 300
Db 241 ORLENDLSGVTLTDEVTYLMDCSFDITSTVDTKLSPPFCDLFTHDWMIOYDLSLG 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHSTNHTLDSNPATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHSTNHTLDSNPATPLNATLYADES 360

QY 361 HNTWVSIFFFALGNGYKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAKEP 420
Db 361 HDNGIISILFALGNGYKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAKEP 420

QY 421 LVRVLVNDRVVPLHCGVDKLGCRKRDFFVGLSFARSGGNWECFA 467
Db 421 LVRVLVNDRVVPLHCGVDKLGCRKRDFFVGLSFARSGGNWECFA 467

RESULT 9
US-08-419-448-32
; Sequence 32, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Seltan
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1988
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448

```

```

FILING DATE: 10-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20026.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-419-448-32

Query Match      75.4%; Score 1862; DB 2; Length 467;
Best Local Similarity 74.5%; Pred. No. 5.2e-188;
Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCFPEITSHLWQYSPFFSLADE 60
Db 1 MGVSAYLLPLLYLLSGVTSLGAVPASRNQSSCDTVDQGYQCFSEITSHLWQYAPFFSLANE 60

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFRASGSDRVIASAEKFEIGFQSA 180
Db 121 YSLGADDLTPFGQGMVNSGKIFRYRYKALARKIVPFRASGSDRVIASAEKFEIGFQSA 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNILDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVINVIPEGAGYNNILDHGLCTAFEESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVDVNLMDMCPEDTIVARTSDATQSPFCDLFTHDWMIOYDLSLG 300
Db 241 ORLENDLSGVTLTDEVTYLMDCSFDITSTVDTKLSPPFCDLFTHDWMIOYDLSLG 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHSTNHTLDSNPATPLNATLYADES 360
Db 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHSTNHTLDSNPATPLNATLYADES 360

QY 361 HNTWVSIFFFALGNGYKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAKEP 420
Db 361 HDNGIISILFALGNGYKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAKEP 420

QY 421 LVRVLVNDRVVPLHCGVDKLGCRKRDFFVGLSFARSGGNWECFA 467
Db 421 LVRVLVNDRVVPLHCGVDKLGCRKRDFFVGLSFARSGGNWECFA 467

RESULT 10
US-08-819-825-3
; Sequence 3, Application US/08819825
; Patent No. 5866118
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: No. 5866118o No. 5866118disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,825
 FILING DATE: 18-MAR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4758-200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-819-825-3

Query Match 75.4%; Score 1862; DB 2; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSGTALGPRGNHSCDVTDDGYQCFPEISHLWGOYSPFFSLADE 60
 Db 1 MGVSAVLLPLLYLLSGVTSGLAVPASRNQSCDVTDDGYQCFSESHLWGOYAPFFSLANE 60

QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
 Db 61 SVISPEVPAGCRVTFQAQLSRHGARYPTDSKGGKYSALIEIEIQONATTFDCKYAFKTYN 120

QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 Db 121 YSLGADDLTPFGEOLVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180

QY 181 KLADPGANPHQASPVINVIPEAGYNTLUDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLKDPRAQPCQSSPKIDVIVISEASSNNTLDPGCTVFDESELADTVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIOYDYLQSLG 300
 Db 241 QRLNDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPFCDLTHDEWINYDYLQSLK 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGYGAGNPLGPTQGVGYANELIARLTHSPVDDDTSSNHTLDSNPATFPLNATLYADFS 360

QY 361 HNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAVEMMQCAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEKEP 420

QY 421 LVRVLNDRVPLHCCPVDAIGRCKRDFVGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHCCPVDAIGRCKRDFVGLSFARSGGNWEECEFA 467

RESULT 11
 US-09-163-642-3
 ; Sequence 3, Application US/09163642
 ; Patent No. 6221644
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Ray, Michael W.
 ; APPLICANT: Klotz, Alan V.
 ; TITLE OF INVENTION: Polypeptides Having Phytase Activity
 ; TITLE OF INVENTION: And Nucleic Acids Encoding Same
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 62216440 No. 6221644disk of No. 6221644th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/163,642
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,825
 FILING DATE: 18-MAR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4758-200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-163-642-3

Query Match 75.4%; Score 1862; DB 4; Length 467;
 Best Local Similarity 74.5%; Pred. No. 5.2e-188;
 Matches 348; Conservative 44; Mismatches 75; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSGTALGPRGNHSCDVTDDGYQCFPEISHLWGOYSPFFSLADE 60
 Db 1 MGVSAVLLPLLYLLSGVTSGLAVPASRNQSCDVTDDGYQCFSESHLWGOYAPFFSLANE 60

QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIOKNATAFKGYAFLKTYN 120
 Db 61 SVISPEVPAGCRVTFQAQLSRHGARYPTDSKGGKYSALIEIEIQONATTFDCKYAFKTYN 120

QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180
 Db 121 YSLGADDLTPFGEOLVNSGKIFRYRYKALARKIVPVRASGSDRVIASAEKIEGFQSA 180

QY 181 KLADPGANPHQASPVINVIPEAGYNTLUDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLKDPRAQPCQSSPKIDVIVISEASSNNTLDPGCTVFDESELADTVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQSPFCDLFTHDEWIOYDYLQSLG 300
 Db 241 QRLNDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPFCDLTHDEWINYDYLQSLK 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGYGAGNPLGPTQGVGYANELIARLTHSPVDDDTSSNHTLDSNPATFPLNATLYADFS 360

QY 361 HNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAVEMMQCAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEKEP 420

QY 421 LVRVLNDRVPLHCCPVDAIGRCKRDFVGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHCCPVDAIGRCKRDFVGLSFARSGGNWEECEFA 467

RESULT 12
 US-09-155-855-3

us-09-488-265-26.ra

Sat Oct 27 15:25:39 2001

```

; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
; US-09-155-855-1

```

```

Query Match 73.4%; Score 1813; DB 4; Length 443;
Best Local Similarity 75.1%; Pred. No. 7.1e-183; Indels 0; Gaps 0;
Matches 331; Conservative 49; Mismatches 61;

QY 27 NSHSCDVTGGYQCFPEJSHLWGOYSPFFSLADESAISPDVFKCVRVTFVQVLSRHGARY 86
Db 3 NQSTCDTVDGQYCFSETHLWGOYAPFFSLANKSAISPDVPAAGCHVTFAQVLSRHGARY 62
QY 87 PTSSKSKYSALIEATOKNATAFKGYAFLKTYNYTLGADDLTPEGEOMVNSGKIFRR 146
Db 63 PTDSKSKYSALIEETQONATTPEGYAFLKTYNYSLGADDLTPEGEOLVNSGVKFYOR 122
QY 147 YKALARKYVFPVFRASGSDRVASAEKFIQFQSAKLADPGANPHQASPVINVIPEGAGY 206
Db 123 YESLTRNIVFPIKSSSSSRVIAKNGKFIQFQSTKLKDPRAQGOSSPKIDVISEASTS 182
QY 207 NNTLDHGLCTAFSESELGDDVEANFTAVFAPTRAPLEAHLPGVNLTDDEVVNLMDMCPF 266
Db 183 NNTLDHGLCTAFSESELGDDVEANFTAVFAPTRAPLEAHLPGVNLTDDEVVNLMDMCPF 242
QY 267 DTVARTSDATQLSPCDLFTHEWIMQYDYLQSLGKYGYGAGNPLGPAQGVGVNELIAR 326
Db 243 DTSTVSTVDTKLSPCDLFTHEWIMQYDYLQSLGKYGYGAGNPLGPAQGVGVNELIAR 302
QY 327 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSHONTMWSIFPAGLYNGTKPLSTTSV 386
Db 303 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSHONTMWSIFPAGLYNGTKPLSTTSV 362
QY 387 ESIEETDCYAAASWVPPFAARAYVEMMOCEAKEPLVRLVNDVRVPLHGGCVDKLGRCKR 446
Db 363 ENITQDTGFSASWVPPFAARAYVEMMOCEAKEPLVRLVNDVRVPLHGGCVDKLGRCKR 422
QY 447 DDFVEGLSFARSGGNWEECPA 467
Db 423 DDFVEGLSFARSGGNWEECPA 443

```

```

RESULT 14
US-09-155-855-2
; Sequence 2, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0

```

```

; Sequence 3, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Aspergillus niger
; US-09-155-855-3

```

```

Query Match 75.1%; Score 1855; DB 4; Length 467;
Best Local Similarity 73.4%; Pred. No. 2.8e-187; Indels 0; Gaps 0;
Matches 343; Conservative 51; Mismatches 73;

QY 1 MGVFVLLSIATLFGSTGALPRGNHSCDVTVDGQYCFPEIHLWGOYSPFFSLADE 60
Db 1 MGVSALLPLVLLSGVTSGLAVPASRQSTCDTVDGQYCFSETHLWGOYAPFFSLANK 60
QY 61 SAISPDVFKCVRVTFVQVLSRHGARYPTSSKSKYSALIEATOKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPAAGCHVTFAQVLSRHGARYPTDSKSKYSALIEETQONATTPEGYAFLKTYN 120
QY 121 YTLGADDLTPEGEOMVNSGKIFRRYKALARKYVFPVFRASGSDRVASAEKFIQFQSA 180
Db 121 YSLGADDLTPEGEOLVNSGVKFYORYESLTRNIVFPIKSSSSSRVIAKNGKFIQFQST 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNTLHGLCTAFSESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQGOSSPKIDVISEASTSNHTLDSNPATFPLNATLYADFS 240
QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDTVARTSDATQLSPCDLFTHEWIMQYDYLQSLG 300
Db 241 ORLENLDSVSLTDTFTVYLMDCSFDTSTSTVDTKLSPCDLFTHEWIMQYDYLQSLN 300
QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGTPQGVGVANELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
QY 361 HDNTMWSIFPAGLYNGTKPLSTTSVETDGYAAASWVPPFAARAYVEMMOCEAKEP 420
Db 361 HDNGIILFALGLYNGTKPLSTTSVETDGYAAASWVPPFAARAYVEMMOCEAKEP 420
QY 421 LVRVLVNDVRVPLHGGCVDKLGRCKRDDFVEGLSFARSGGNWEECPA 467
Db 421 LVRVLVNDVRVPLHGGCVDKLGRCKRDDFVEGLSFARSGGNWEECPA 467

```

```

RESULT 13
US-09-155-855-1
; Sequence 1, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

```

; SEQ ID NO 2

; LENGTH: 443

; TYPE: PRT

; ORGANISM: Aspergillus niger

; FEATURE:

; OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown

US-09-155-855-2

Query Match

Best Local Similarity 73.2%; Score 1808; DB 4; Length 443;

Matches 330; Conservative 49; Mismatches 59; Indels 0; Gaps 0;

QY 30 SCDTVDGYGOCFPEISHLWGQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTS 89

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 TCDTVGQYOCFSETSHLWGQYAPFFSLANKSAISPDVAGCHVTFPAQVLSRHGARYPTD 65

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 90 SKGKYSALIEAQKNATPKYAFKTYNYTLGADDLTPFGEQOMVNSGKFKYRKYKA 149

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 66 SKGKYSALIEEQNATTFEGKYAFKTYNSLGADDLTPFGEQELVNSGVKFKYQYVES 125

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 150 LARKIVPFRASGSDRVIAAEKFIKGFOSAKIADPGANPHOASPVINVIIPGAGYNN 209

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 126 LTRNIIVPFRSSGSSRVIASGNKFIKGFQSTKLDPRAQPGOSSPKRIDVISEASTSN 185

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 210 LDHGLCTAFEESELDGDDVANTAVFAPPIRARLEAHLPGVNLTDDEVVNLDMCPFD 269

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 186 LDPGCTVFESELDADTEANFATFVPSIRORLENDLSGVSGLTDEVTYLMDCSFD 245

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 270 ARTSDATOLSPCEDLFTHDEWIOYDLOSLGKYGYGAGNPLGPAQGVGFVNELIAR 329

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 246 STSTVDTKLSPCEDLFTHEWINYDLOSLKYYGHGAGNPLGPTQGVGYANELIAR 305

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 330 SPVQDHTSNHTLDSNPATFPLNATLYADFSDHNTMVSIFPAGLYNGTKPLSTTS 389

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 306 SPVHDDTSNHTLDSNPATFPLNSTLYADFSDHNGIISILFALGLYNGTKPLST 365

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 330 EETDGYAASWTVPFAARAYVEMOCEAKEPLRVLVNDRVVPVPLHGCVDKLGCRK 449

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 366 TQDGFSSAWTVFASRMVYVEMOCSOEPLRVLVNDRVVPVPLHGCVDALGRCTR 425

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 450 VEGLSFARSGGNWEECEFA 467

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 426 VKGLSFARSGGDWGECEFA 443

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

RESULT 15

US-08-819-825-2

; Sequence 2, Application us/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5866118 of No. 5866118 disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 4758.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 867 0123
 TELEFAX: 212 867 0298
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 475 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-819-825-2

Query Match 56.5%; Score 1395; DB 2; Length 475;
 Best Local Similarity 58.5%; Pred. No. 1.1e-138;
 Matches 285; Conservative 59; Mismatches 105; Indels 38; Gaps 11;

QY 1 MGVP-VVLLSTATLFGSTSGTALGP--RGNSHGCDTVDDGGYQCFPEISHLWGQYSPFSL 57

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 LGSLVLVLLOFSALL--TASPAIPPFWRKHPNVD-----IARHWGQYSPFSL 52

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 58 ADESAISPDVPGCRVTFVQVLSRHGARYPTS SSKSKYSALIEAOKNATFAKGYAFLK 117

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 53 AEVSEISPAVPKGRVTFVQVLSRHGARYPTAHKSEYVYAEIQLRIQDTATEFKGDEAFLR 112

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 118 TNYTLGADDLTPFGEQOMVNSGKFKYRKYKALARKIVPFRASGSDRVIASAEKIEGF 177

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 113 DYAYHLGADNLTRFGEQOMESGRQFVHYRREOAREIVPFRAGSARVIASAEFFNRGF 172

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 178 QSAKLADPGANPHOASPVINVIIPGAGYNTLDHGLCTAFEESELDGDDVE--ANFTAVFA 236

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 173 QDAKDRDRPSNKDOAEFVINVIISSETSGNNTLDGLTCPAAEAP--DPTQPAEFLQVFG 230

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 237 PPIRABLEAHLPGVNLTDDEVVNLDMCPFDVARTSDAT----OLSPCEDLFTHDEWIO 292

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 231 PRVKKITKHPGVNLTDDEVVNLDMCPFDVARTSDAT----OLSPCEDLFTHDEWIO 288

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 293 YDYLTSLGKYGYGAGNPLGPAQGVGFVNELIARLT--HSPVQDHTSNHTLDSNPATEPL 351

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 289 YDYLTSLGKYGYGAGNPLGPAQGVGFVNELIARLT--HSPVQDHTSNHTLDSNPATEPL 348

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 352 NATLYADFSDHNTMVSIFPAGLYNGTKPLSTTSVE--SIBETDGYAASWTVPFAARAYV 409

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 349 DAVLYADFSDHNTMVSIFPAGLYNGTKPLSTTSVE--SIBETDGYAASWTVPFAARAYV 408

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 410 ENMOC-----EAEKEPLRVLVNDRVVPVPLHGCVDKLGCRKRDDEVEGLSFARSQ 459

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 409 ELIRCETETSEEEEGEDEPFVRLVNDVRVPLHGCVRDRWGRCDRDEWIKGLTFARQ 468

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 460 GNWEECF 466

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 469 GHWDRCF 475

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

Search completed: October 26, 2001, 16:40:45
 Job time: 4956 sec

Sat Oct 27 15:25:39 2001

us-09-488-265-26.rai

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:29 ; Search time 57.24 Seconds
(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGFEVLLSLIATLFGSTSGT.....DFVEGLSPARSGGNWEECF 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2470	100.0	467	21	AB20524
2	2470	100.0	467	21	AA69566
3	2448	99.1	467	20	AA43170
4	2448	99.1	467	21	AA20527
5	2448	99.1	467	21	AA69569
6	2442	98.9	467	21	AA20533
7	2436	98.6	467	21	AB20534
8	2353	95.3	467	21	AA20531
9	2347	95.0	467	21	AA20532
10	2344	94.9	441	21	AA20523
11	2342	94.8	467	20	AA43169

12	2342	94.8	467	21	AB20526	Consensus phytase
13	2342	94.8	467	21	AA69568	Mutant phytase-1,
14	2311	93.6	467	20	AA93380	Fungal phytase pro
15	2311	93.6	467	21	AA20515	Consensus phytase
16	2311	93.6	467	21	AA69558	Phytase-1, a conse
17	2305	93.3	467	20	AA69558	Fungal consensus p
18	2304	93.3	467	20	AA93382	Fungal consensus p
19	2304	93.3	467	20	AA93383	Fungal consensus p
20	2296	93.2	467	20	AA93906	Ascomycete consens
21	2296	92.9	467	20	AA93384	Fungal consensus.p
22	2295	92.9	467	20	AA93385	Fungal consensus p
23	2291	92.8	467	21	AA20530	Consensus phytase
24	2291	92.8	467	21	AA20530	Consensus phytase
25	2204	89.2	467	20	AA69572	Conspyl12, a deriv
26	2204	89.2	467	20	AA43172	Phytase-7, a deriv
27	2201	89.1	467	21	AA69571	Consensus phytase
28	2185	88.5	441	21	AB20514	Consensus phytase
29	2115	85.6	437	21	AB20525	Phytase-11, a cons
30	2115	85.6	437	21	AA69567	Consensus phytase
31	2066	83.6	431	21	AA20535	Consensus phytase
32	2066	83.6	431	21	AA69557	Initial consensus
33	2050.5	83.0	424	21	AA20536	Consensus phytase
34	2050.5	83.0	424	21	AA69565	Consensus phytase
35	2048	82.9	467	21	AA69574	Initial consensus
36	2028	82.1	467	20	AA69574	Mutant Aspergillus
37	2028	82.1	467	21	AA43171	A. fumigatus phyt
38	2028	82.1	467	21	AA20528	Aspergillus fumiga
39	1902	77.0	467	21	AA69570	Mutant Aspergillus
40	1902	77.0	465	19	AA84356	Aspergillus fumiga
41	1892.5	76.6	440	21	AA69549	A. fumigatus phyt
42	1889.5	76.5	440	21	AA20507	Aspergillus fumiga
43	1889.5	76.5	440	21	AA69550	Aspergillus fumiga
44	1884.5	76.3	440	21	AA20508	Aspergillus fumiga
45	1884.5	76.3	440	21	AA69551	Aspergillus fumiga

ALIGNMENTS

RESULT 1
AAB20524
ID AAB20524 standard; Protein: 467 AA.
XX
AC AAB20524;
XX
DT 05-DEC-2000 (first entry)
XX
DE Consensus phytase 10 SEQ ID NO:26.
XX
KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure.
XX
OS Synthetic.
XX
PN WO200043503-A1.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-DK00025.
XX
PR 22-JAN-1999; 99DK-0000092.
PR 21-SEP-1999; 99DK-0001340.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Lehmann M;
XX
DR WPI; 2000-491161/43.
DR N-PSDB; AAA73232.
XX

PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 1; Fig 5a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 2470; DB 21; Length 467;

Best Local Similarity 100.0%; Pred. No. 3.5e-240; Indels 0; Gaps 0;
 Matches 467; Conservative 0; Mismatches 0;

QY 1 MGVFVLLSLATLFGSTGALPGNSHSCDFTVDGGYQCFPEISHLWGQYSPFFSLADE 50
 DB 1 mgfvvllslatlfgstgaltprgnshscdtvddggyqcfpeishlwggyspffslade 60
 QY 61 SAISPDVPGKCRVTFVOVLGRHGRVPTSSKKYSALIEATQKNATAPKGYAFLKTYN 120
 DB 61 saispdvpgkcrvtfvqlsrhgrvptsskkskysalieleatqknatapkgyaflkty 120
 QY 121 YTLGADDLTPGEOOMVNSGKIFRYRKALARKIVPFRASGSDRVITASAERKIEGFQSA 180
 DB 121 ytlgaddltpfgeodmvsngkifryrkalkarkivpfrasgsdrvitasaekefiesq 180
 QY 181 KLADPGANPHQASPVNVIPEGAGYNNTLDHGLCTAFESSELGDDVEANFTAVEAPPTR 240
 DB 181 kladpganphqaspvniipegagynntldhglctafeeseligddveanftavappir 240
 QY 241 ARLEAHLPGVNLTDYVNLMDMCPFTDVARTSDATLSLSPFCDLFTHDEWIQYDYLQSLG 300
 DB 241 arleahlpgvnltdyvnldmcpfdtvdartsdatslspfcdlftdhewiqdyylqslg 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPPLNATLYADFS 360
 DB 301 kygygagnplgpaggvfnelliarlthspvqdhstnhtldsnpatfpplnatlyadfs 360
 QY 361 HDNTMVSITFFALGLYNGTKPLSTTSVESIETDCYAAASWTVPFAARAYVEMMQCEAEKEP 420
 DB 361 hdntmvsiffalglyngtkplsttsvesieetdcyaaaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHCGGVKLGKCRKDDFVEGLSFARSGGNWCEFA 467
 DB 421 lvrvlndrvplhcggvkldgkcrddfveglslfarsgggnwcecf 467

RESULT 2

AAV69566

ID AAV69566 standard; protein: 467 AA.

XX AC AAV69566;

XX DT 19-APR-2000 (first entry)

XX DE Phytase-10, a consensus phytase.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytase-10.
 KW phosphate availability; consensus; phytase-10.

XX Aspergillus terreus 9A1.
 OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger str. NRRL3135.
 OS Aspergillus fumigatus ATCC13073.
 OS Aspergillus fumigatus ATCC32722.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Emericella nidulans.
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Paxillus involutus NN005693.
 OS Trametes pubescens NN9343.
 OS Agrocybe pediades NN009289.
 OS Peniophora lycii NN006113.
 OS Thermomyces lanuginosa.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Peptide 1..26
 FT /note= "Phytase signal peptide from Aspergillus terreus
 FT cbs16.46"
 FT Protein 27..467
 FT /note= "Mature phytase-10 consensus"
 XX EP969089-A1.
 XX 05-JAN-2000.
 XX 23-JUN-1999; 99EP-0111949.
 XX 29-JUN-1998; 98EP-0111960.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX Bruggen R, Lehmann M, Wyss M;
 WPI; 2000-099429/09.
 N-PSDB; AAZ59642.
 PT New stabilised enzyme formulation, useful for feed compositions for
 monogastric animals -
 XX Example 4; Fig 17; 101pp; English.
 CC The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion,
 CC and pelleting. The present sequence represents a consensus sequence,
 CC designated phytase-10, which was derived from the mature phytase
 CC sequences from a variety of fungi (AAV69544-Y69546, AAV69548-Y69556,
 CC AAV69564) and the Basidiomycetes phytase consensus AAV69563 and
 CC additionally contains the Aspergillus terreus cbs16.46 signal peptide at
 CC the N-terminus.


```

XX SQ Sequence 467 AA;
Query Match 100.0%; Score 2470; DB 21; Length 467;
Best Local Similarity 100.0%; Pred. No. 3.5e-240;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCPEISHLWGOYSPFFSLADE 60
Db 1 mgvfvlisiatlfgstgctalgrgnshscdtvdggycipeishlwggyspffslade 60
Qy 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 saispdvpkgrvtfvqvlshrgharyptsskkskysalialieaiqknatafkgyaflkty 120
Qy 121 YTLGADDLTPGEOQWNSGKIFVRRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltpgeqmwngskifvrrykalarkivpfvrsgsdrviasaekfieqfsgsa 180
Qy 181 KLADPGANPHOASPVINVIIEGAGYNNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 kladpganphoaspvinviiepegagynntldhglctafeeselgddveanftavfappir 240
Qy 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPCDLFTHDEWIQDYQLQSLG 300
Db 241 arleahlpvnltdedvvnldmcpfdvtartsdatsqlspcdlftdhewiqdyqlqslg 300
Qy 301 KYYGYGAGNPLGPAQGVGVNLIARLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 kyygygagnplgpaqgvgnvliarlthspvqdhstnnhtldsnpatfplnatlyadfs 360
Qy 361 HDNTMVSIFFFALGYNLTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 hdntmvsiffalgyngtkplsttsvesieetdgyaaswtvpfaarayvemmqceakep 420
Qy 421 LVRVLVNDVRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWECFA 467
Db 421 lvrvlvndrvvplhcgvdkgkrdrdfveglsfarsggnwecfa 467

RESULT 3
AAY43170
ID AAY43170 standard; Protein; 467 AA.
XX AC AAY43170;
XX DT 06-JAN-2000 (first entry)
XX DE Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.
XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
XX KW consensus sequence.
XX OS Synthetic.
XX PN WO9948380-A1.
XX PD 30-SEP-1999.
XX PF 22-MAR-1999; 99WO-DK00154.
XX PR 23-MAR-1998; 98DK-0000407.
XX PR 19-JUN-1998; 98DK-0000806.
XX PR 18-SEP-1998; 98DK-0001176.
XX PR 22-JAN-1999; 99DK-0000091.
XX PR 22-JAN-1999; 99DK-0000093.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Petersen S;
XX PX WPI; 1999-591030/50.
XX OS

N-PSDB; AAZ31521.
Preparing animal feed using a thermostable phytase -
Example 3; Fig 10; 71pp; English.
This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
The invention relates to a process for preparing animal feed by
agglomerating feed ingredients with a thermostable phytase, which is
added before or during agglomeration. The thermostable phytase is useful
for expression in transgenic plants. These plants are useful in the
preparation of animal feed itself. The thermostable phytase allows animal
feed to be produced more efficiently, in addition to improved
phytase-expressing transgenic plants. These plants provide a feed
ingredient and a feed additive (phytase) simultaneously.

SQ Sequence 467 AA;
Query Match 99.1%; Score 2448; DB 20; Length 467;
Best Local Similarity 98.9%; Pred. No. 5.9e-238;
Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCPEISHLWGOYSPFFSLADE 60
Db 1 mgvfvlisiatlfgstgctalgrgnshscdtvdggycipeishlwgtyspffslade 60
Qy 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 saispdvpkgrvtfvqvlshrgharyptsskkskysalialieaiqknatafkgyaflkty 120
Qy 121 YTLGADDLTPGEOQWNSGKIFVRRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltpgeqmwngskifvrrykalarkivpfvrsgsdrviasaekfieqfsgsa 180
Qy 181 KLADPGANPHOASPVINVIIEGAGYNNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 kladpganphoaspvinviiepegagynntldhglctafeeselgddveanftavfappir 240
Qy 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATQLSPCDLFTHDEWIQDYQLQSLG 300
Db 241 arleahlpvnltdedvvnldmcpfdvtartsdatsqlspcdlftdhewiqdyqlqslg 300
Qy 301 KYYGYGAGNPLGPAQGVGVNLIARLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
Db 301 kyygygagnplgpaqgvgnvliarlthspvqdhstnnhtldsnpatfplnatlyadfs 360
Qy 361 HDNTMVSIFFFALGYNLTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMQCEAKEP 420
Db 361 hdntmvsiffalgyngtkplsttsvesieetdgyaaswtvpfaarayvemmqceakep 420
Qy 421 LVRVLVNDVRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWECFA 467
Db 421 lvrvlvndrvvplhcgvdkgkrdrdfveglsfarsggnwecfa 467

RESULT 4
AAB20527
ID AAB20527 standard; Protein; 467 AA.
XX AC AAB20527;
XX DT 05-DEC-2000 (first entry)
XX DE Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.
XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
XX KW temperature stability; pH profile; temperature profile; reaction rate;
XX KW specific activity; substrate specificity; substrate cleavage pattern;
XX KW substrate binding; position specificity; phytate degradation rate;
XX KW food; feed; phytate; manure.
XX OS Synthetic.

```


CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-10
 CC consensus sequence, phytase-10-thermo[31]-Q50T-K91A, which has a
 CC temperature optimum and melting point 4 degrees Celsius higher than that
 CC of phytase-10 (AAY69566). Its specific activity with phytate as a
 CC substrate is also strongly increased.

XX Sequence 467 AA;

Query Match 99.1%; Score 2448; DB 21; Length 467;
 Best Local Similarity 98.9%; Pred. No. 5.9e-238;
 Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVPVLLSTATLFGSTGALGRGNSHSDTVGQYQCFPEISHLMGQYSPFFSLADE 60
 DB 1 mgvfvllsatiifgsgtalgrgnschsdtdvgyqcfpeishlwgtyspffslade 60
 QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 saispdpvkgcrvtfvqvlshrgharyptsskksalsaliedaiaqknatafkgyaflkty 120
 QY 121 YTLGADDLTPGEGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrryrykalkarkivpfiragsdrviasaekfiqfsga 180
 QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESLGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinviiepegagynntldhglctafeeslgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSPPCDLTFHDEWIOYDYLQSLG 300
 DB 241 arleahlpgvnltdvnlmdmcpdftvartsdatsqlspcdlftdhewiogydylqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnpplgpaqgvfveliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyasaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRKDDFVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlvndrvvplhgcvdklgcrkddfveglfsfarsggnwecfa 467

RESULT 6

AAB20533
 ID AAB20533 standard; Protein; 467 AA.

AC AAB20533;

XX 05-DEC-2000 (first entry)

DE Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95.

KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.

XX WO200043503-A1.
 PN 27-JUL-2000.
 PD 21-JAN-2000; 2000WO-DK00025.
 PF 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PI Lehmann M;
 DR WPI: 2000-491161/43.
 DR N-PSDB; AAN73292.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 PS Disclosure; Fig 24a-c; 240pp; English.
 XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, the velocity and level of
 CC substrate binding, position specificity, the reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX Sequence 467 AA;

Query Match 98.9%; Score 2442; DB 21; Length 467;
 Best Local Similarity 98.7%; Pred. No. 2.4e-237;
 Matches 461; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGVPVLLSTATLFGSTGALGRGNSHSDTVGQYQCFPEISHLMGQYSPFFSLADE 60
 DB 1 mgvfvllsatiifgsgtalgrgnschsdtdvgyqcfpeishlwgtyspffslade 60
 QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFLKTYN 120
 DB 61 saispdpvkgcrvtfvqvlshrgharyptsskksalsaliedaiaqknatafkgyaflkty 120
 QY 121 YTLGADDLTPGEGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgeqmvnsgikfrryrykalkarkivpfiragsdrviasaekfiqfsga 180
 QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESLGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinviiepegagynntldhglctafeestlgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSPPCDLTFHDEWIOYDYLQSLG 300
 DB 241 arleahlpgvnltdvnlmdmcpdftvartsdatsqlspcdlftdhewiogydylqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnpplgpaqgvfveliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyasaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRKDDFVEGLSFARSGGNWEECFA 467

Db 421 lvrvlndrvplhgcavdklgrckrdffveglstfarsgnweecfa 467
 |||

RESULT 7
 AAB20534
 ID AAB20534 standard; Protein; 467 AA.

XX AAB20534;
 XX
 DT 05-DEC-2000 (first entry)

DE Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73293.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 25a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 98.6%; Score 2436; DB 21; Length 467;
 Best Local Similarity 98.5%; Pred. No. 9.5e-237;
 Matches 460; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MGVFVLLSIALFGSTGALPGRNHSCDVTVDGGYOCFPEISHLWGYSPFFSLADE 60

Db 1 mgvfvlslatlfsgtsalgrgnshscdvtvdggyqcfpeishlwgtyspffslade 60
 |||

QY 61 SALSPPDPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAQKNATAPKGYAFUKTYN 120

Db 61 salspdpkgcrvtfvqvlshrgharyptsskaskysallieaqlgnatarkgyafuktyn 120
 |||

QY 121 YTLGADDLTPFGEQQVMNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 |||
 Db 121 ytlgaddltpfgeqmqvmsgkfyrrykalarikiyfriragsdrviasaekfiegqsa 180
 |||
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTATPESELGDDVEANFTAVFAPPIR 240
 |||
 Db 181 kladpganphgaspvinvipegagynntldhglctateestlgddveanftavfappir 240
 |||
 QY 241 ARLEAHLPGVNLTDDEVDVNLDMCPDFTVARTSDATQLSPPCDLFTHDEWIOYDYLSLGL 300
 |||
 Db 241 arleahlpgvnltdedvnlmdmcpdftvartsdatsqlspcdlftdhewiodylslgl 300
 |||
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
 |||
 Db 301 kyygygagnplgagvgvfneliarlthspvgdhtstnhtldsnpatfplnatlyadfs 360
 |||
 QY 361 HNTWVSFFALGLNGYKPLSTTSVESIEEDDGAASWTVPFAARAYVEMMQCAEKPEP 420
 |||
 Db 361 hntwmvsiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcaekpep 420
 |||
 QY 421 LVRVLVNDRVPLHGGGVNDKGRCKRDDFVEGLSFARSGNWECCFA 467
 |||
 Db 421 lvrvlndrvplhgcavdklgrckrdffveglstfarsgnweecfa 467
 |||

RESULT 8

AAB20531

ID AAB20531 standard; Protein; 467 AA.

AC AAB20531;

XX 05-DEC-2000 (first entry)

Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce

CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.3%; Score 2353; DB 21; Length 467;
 Best Local Similarity 94.9%; Pred. No. 2.2e-228;
 Matches 443; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRGNSHSCDVTGQYQCPEISHLWGOYSPFFSLADE 60
 Db 1 mgvfvllsstatlfgstsgtalgrgnschscdtvggyqcfpeishlwgtyspyfslade 60

QY 61 SAISPDVPKCRVTFVQVLSRHGARYPTSSKSKYALIEAIOKNAFAFKGYAFLKTYN 120
 Db 61 saispdvpkdcrtvfvqlsrhgaryptsakskaysaliaeqknatafkgyaflkty 120

QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPVRASGSDRVITASAEKFEQFSA 180
 Db 121 ytlgaddltpfgenqmvnsqikfrryrykalkarkivpfiragsdrviasaeekflegf 180

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpgsqphqaspvinnviiepgsgynntldhglctafedstlgddveanftalfap 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPEDTARTSDATQLSFCDLFTHDEWIQDYLOSLG 300
 Db 241 arleahlpgvnltdedvvylnmdmcpdrtartsdatslfcalfthdewiqdyloqslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNLIELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 Db 301 kygygagpnlgpaqvggfaneliarlthspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdntmsiffalglyngtkplsttsvesieetdgyaaswtvpfaarayvemmqcqaek 420

QY 421 LVRVLNDRVPLHGGVVDKLGRCRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 lvrvlndrvvplhggcvdklgrckrddfvglsfarsggnwaecfa 467

RESULT 9

AAB20532
 ID AAB20532 standard; Protein; 467 AA.

XX AC AAB20532;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 XX temperature stability; pH profile; temperature profile; reaction rate;
 XX specific activity; substrate specificity; substrate cleavage pattern;
 XX substrate binding; position specificity; phytate degradation rate;
 XX food; feed; phytate; manure.

XX OS Synthetic.

XX PN W0200043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX XX

PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX DR WPI; 2000-491161/43.

XX DR N-PSDB; AAA73291.

XX PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 XX compound foods and feeds -

XX PS Disclosure; Fig 23a-c; 240pp; English.

XX CC The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX SQ Sequence 467 AA;

Query Match 95.0%; Score 2347; DB 21; Length 467;
 Best Local Similarity 94.8%; Pred. No. 8.9e-228;
 Matches 442; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRGNSHSCDVTGQYQCPEISHLWGOYSPFFSLADE 60
 Db 1 mgvfvllsstatlfgstsgtalgrgnschscdtvggyqcfpeishlwgtyspyfslade 60

QY 61 SAISPDVPKCRVTFVQVLSRHGARYPTSSKSKYALIEAIOKNAFAFKGYAFLKTYN 120
 Db 61 saispdvpkdcrtvfvqlsrhgaryptsakskaysaliaeqknatafkgyaflkty 120

QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPVRASGSDRVITASAEKFEQFSA 180
 Db 121 ytlgaddltpfgenqmvnsqikfrryrykalkarkivpfiragsdrviasaeekflegf 180

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpgsqphqaspvinnviiepgsgynntldhglctafedstlgddveanftalfap 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPEDTARTSDATQLSFCDLFTHDEWIQDYLOSLG 300
 Db 241 arleahlpgvnltdedvvylnmdmcpdrtartsdatslfcalfthdewiqdyloqslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNLIELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 Db 301 kygygagpnlgpaqvggfaneliarlthspvqdhstnhtlidsnpatfplnatlyadfs 360

QY 361 HDNTWVSIFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdntmsiffalglyngtkplsttsvesieetdgyaaswtvpfaarayvemmqcqaek 420

QY 421 LVRVLNDRVPLHGGVVDKLGRCRDDFVEGLSFARSGGNWEECEFA 467
 Db 421 lvrvlndrvvplhggcvdklgrckrddfvglsfarsggnwaecfa 467

RESULT 10

AAB20523

ID AAB20523 standard; Protein; 441 AA.

XX AC AAB20523;

XX XX

DT	05-DEC-2000 (first entry)	
XX	Consensus phytase 10 (Fcp10) SEQ ID NO:24.	
DE	Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;	
XX	temperature stability; pH profile; temperature profile; reaction rate;	
KW	specific activity; substrate specificity; substrate cleavage pattern;	
KW	substrate binding; position specificity; phytate degradation rate;	
KW	food; feed; phytate; manure.	
XX	Synthetic.	
OS		
XX	WO200043503-A1.	
PN	27-JUL-2000.	
XX		
XX	21-JAN-2000; 2000WO-DK00025.	
PF		
XX	22-JAN-1999; 99DK-0000092.	
PR		
PR	21-SEP-1999; 99DK-0001340.	
XX		
XX	(NOVO) NOVO NORDISK AS.	
PA		
XX	Lehmann M;	
PI		
XX	WPI; 2000-491161/43.	
DR		
XX	Novel phytases with improved properties such as temperature stability,	
PT	pH stability and substrate specificity, for use in pharmaceuticals and	
PT	compound foods and feeds -	
PT		
XX	Example 2; Fig 4a-d; 240pp; English.	
PS		
XX	The present invention describes improved phytases, preferably with	
CC	increased thermostability, and methods for producing them. The methods	
CC	can be used for producing phytases with improved properties e.g.	
CC	temperature stability, pH stability, pH profile, temperature profile,	
CC	specific activity, substrate specificity, substrate cleavage pattern,	
CC	substrate binding, position specificity, the velocity and level of	
CC	release of phosphate from corn, reaction rate, phytate degradation rate,	
CC	and end level of released phosphate. The phytases can be used to produce	
CC	pharmaceutical compositions or compound food or feeds. The feed can be	
CC	used to reduce levels of phytate in animal manure, by converting it	
CC	into lower inositol phosphates and/or inositol and inorganic phosphate.	
CC	The present sequence represents a phytase sequence from the present	
CC	invention.	
XX		
XX	Sequence 441 AA;	
XX		
XX	Query Match 94.9%; Score 2344; DB 21; Length 441;	
XX	Best Local Similarity 100.0%; Pred. No. 1.6e-227;	
XX	Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
XX		
QY	27 NNSCSDTVGGYQCFEISHLWQYSPFFSLADESAISPDVPRGCRVTFVQVLSRHGARY 86	
Db	1 nshscdtdggyqcfeshlwgyspffsladesaispdvprgcrvtfvqlsrhgary 60	
QY	87 PTKSKKKYSALIEATQKNATAFKGYAFKTYNTLIGADLLTPFGQQMVSIGKFYRR 146	
Db	61 ptkskkksalieleatqknatafkgyafktyntligadlltpfgeqqmvsigkfyrr 120	
QY	147 YKALARKIVPFRASGSDRVIASAEKFIQFQSAKLADPGANPRQASPVINVIPEGAGY 206	
Db	121 ykalarxivpfrasgsdrviasaekfiefqsaqladpganpqaspvinviipegagy 180	
QY	207 NNTLDHGLCTAFEESELGDDVEANFTAVFAPPTARLEAHLPGVNLTDDEVNLMDCPF 266	
Db	181 nntldhglctafeeselgddveanftavfappirarleahlpvnltdedvnlmdmcpf 240	
QY	267 DTVARTSDATQLSPFCDFLTHDEWIOYDYQLSLGKYGYGAGNPLGPAQGVGFVNELIAR 326	
Db	241 dtvartsdattqlspfcdfldhewigydyqlslgkygygagnplgpaqgvgfvnelliar 300	

QY	327 LTHSPVDHTSTNHTLDGNPATFFPLNATLYADFSDNTMWSIFFALGLYNGFKPLSTTSV 386	
Db	301 lthspvqdhstnhtldnpatffplnatlyadfsdntmwsiffalglngtkplsttsv 360	
QY	387 ESTEETDGYAASWTVPFAARAYVEMMOCEAEKEPLVRVLVNDRVVPLHGCGVDKLGRCR 446	
Db	361 esleetdgyaaswtvpfaarayvemmqeaekeplrvlndrvvplhgcgvdklgrckr 420	
QY	447 DDFVEGLSFARSGGNWECEFA 467	
Db	421 ddfveglsfarsggnweecfa 441	
XX	RESULT 11	
XX	AA43169	
ID	AA43169 standard; Protein; 467 AA.	
XX		
AC	AA43169;	
XX		
DT	06-JAN-2000 (first entry)	
XX		
DE	Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.	
XX		
KW	Phytase; animal feed preparation; thermostable phytase; transgenic plant;	
KW	consensus sequence.	
OS	Synthetic.	
XX		
PN	WO9948380-A1.	
XX		
PD	30-SEP-1999.	
XX		
PF	22-MAR-1999; 99WO-DK00154.	
XX		
PR	23-MAR-1998; 98DK-0000407.	
PR	19-JUN-1998; 98DK-0000806.	
PR	18-SEP-1998; 98DK-0001176.	
PR	22-JAN-1999; 99DK-0000091.	
PR	22-JAN-1999; 99DK-0000093.	
XX		
PA	(NOVO) NOVO-NORDISK AS.	
XX		
PI	Petersen S;	
XX		
DR	WPI; 1999-591030/50.	
DR	N-PSDB; AA231520.	
XX		
PT	Preparing animal feed using a thermostable phytase	
XX		
PS	Example 3; Fig 9; 71pp; English.	
XX		
CC	This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.	
CC	The invention relates to a process for preparing animal feed by	
CC	agglomerating feed ingredients with a thermostable phytase, which is	
CC	added before or during agglomeration. The thermostable phytase is useful	
CC	for expression in transgenic plants. These plants are useful in the	
CC	preparation of animal feed itself. The thermostable phytase allows animal	
CC	feed to be produced more efficiently, in addition to improved	
CC	phytase-expressing transgenic plants. These plants provide a feed	
CC	ingredient and a feed additive (phytase) simultaneously.	
XX		
XX	Sequence 467 AA;	
XX		
XX	Query Match 94.8%; Score 2342; DB 20; Length 467;	
XX	Best Local Similarity 94.4%; Pred. No. 2.8e-227;	
XX	Matches 441; Conservative 10; Mismatches 16; Indels 0; Gaps 0;	
XX		
QY	1 MGVFVLLSTATLFGSTGALPRGNSHSCDNDVGGYQCFEISHLWGOYSPFFSLADE 60	
Db	1 mgvfvlslatlfgstsgalprgnshscdndvgyqcfepelshlwgtyspysfslade 60	

QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAEIKTYN 120
 Db 61 saispdvpddcrvtvqvlshrgharyptssaskaysalieaiknatarkgyafiktyN 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfiragsdrviasaekfieqfsga 180
 QY 181 KLADPGANPHOASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVEANTVAFAPPIR 240
 Db 181 kladpgsghaspsvinviiepgsgyntldhgtctafedseigddveantafafair 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATLSLSPFCDLTFHDEWIQDYLOSLG 300
 Db 241 arleadlpgvnltdedvvyldmcpfdvtartsdatselspfcalfthdewiqdyloqslg 300
 QY 301 KYCYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyygygagnplgpaqgvgfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdnmtisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlvndrvplhgcavdklgrckrdrdfveglisfarsggnwaecfa 467

RESULT 12

AAB20526
 ID AAB20526 standard; Protein; 467 AA.

XX AAB20526;

DT 05-DEC-2000 (first entry)

DE Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; reaction rate;
 KW substrate binding; position specificity; substrate cleavage pattern;
 KW food; feed; phytate; manure.

OS Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,

CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 94.8%; Score 2342; DB 21; Length 467;
 Best Local Similarity 94.4%; Pred. No. 2.8e-227;

Matches 441; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLPGSTGCTALGPRGNSHSCDTVDGGYQCPEISHLMGQYSPFFSLADE 60
 Db 1 mgvfvllsialtlfgstgctalgrgnshscdtvdggycpfeishlmwgtyspyfslade 60
 QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSKSKYSALIEAIQKNATAFKGYAEIKTYN 120
 Db 61 saispdvpddcrvtvqvlshrgharyptssaskaysalieaiknatarkgyafiktyN 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfiragsdrviasaekfieqfsga 180
 QY 181 KLADPGANPHOASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVEANTVAFAPPIR 240
 Db 181 kladpgsghaspsvinviiepgsgyntldhgtctafedseigddveantafafair 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATLSLSPFCDLTFHDEWIQDYLOSLG 300
 Db 241 arleadlpgvnltdedvvyldmcpfdvtartsdatselspfcalfthdewiqdyloqslg 300
 QY 301 KYCYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyygygagnplgpaqgvgfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMVSIFFALGLYNGTKPLSTTSVESIEETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 Db 361 hdnmtisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqcqaekp 420
 QY 421 LVRVLVNDVRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlvndrvplhgcavdklgrckrdrdfveglisfarsggnwaecfa 467

RESULT 13

AAY69568

ID AAY69568 standard; protein; 467 AA.

XX AAY69568;

XX 19-APR-2000 (first entry)

DT Mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.

DE Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutin.

XX Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS Aspergillus niger var. awamori.

OS Aspergillus niger T213.

OS Aspergillus niger str. NRRL3135.

OS Aspergillus fumigatus ATCC13073.

OS Aspergillus fumigatus ATCC32722.

OS Aspergillus fumigatus ATCC58128.

OS Aspergillus fumigatus ATCC26906.

Query Match 93.6%; Score 2311; DB 20; Length 467;
 Best Local Similarity 93.1%; Pred. No. 3.8e-224;
 Matches 435; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGATLPGRNHSCDVTGQYQCFPEISHLWGQYSPFFSLADE 60
 DB 1 mgvfvlisiatlfgstgatalprgnshscdvtggyqcfpeishlwgqyspyfslade 60
 QY 61 SAISPDVPKGRVTFVQLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKTYN 120
 DB 61 saispdvpdkgrvtfvqlsrhgaryptsskksaysalieaigknatafkgyafikty 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTGQFSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfirasgsdrviasaekfiegfsa 180
 QY 181 KLADPGANPHQASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
 DB 181 kladpganphqaspvdiviiepegsgyntldhgtctafedselgddveanftalfapair 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDCPEDTAVRTSDATOLSPFCDLTHDEWTOYDYLSLG 300
 DB 241 arleahlpgvnltdedvvylnmdcpfetavrtsdatslpcalfthdewrqdyylsg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMYSIFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARYVEMMOCEAKEP 420
 DB 361 hdnsmisifalglngtaptlsttsvesieetdgyaswtvpfgarayvemmqcqaek 420
 QY 421 LVRVLVNDVRVPLHGGVCKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlvndrvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

RESULT 15
 AAB20515
 ID AAB20515 standard; Protein; 467 AA.
 AC AAB20515;
 XX XX
 DT 05-DEC-2000 (first entry)
 XX XX
 DE Consensus phytase SEQ ID NO:16.
 XX XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX OS
 OS Synthetic.
 XX OS
 PN WO200043503-A1.
 XX PN
 PD 27-JUL-2000.
 XX XX
 PF 21-JAN-2000; 2000WO-DK00025.
 XX XX
 PR 22-JAN-1999; 93DK-0000092.
 PR 21-SEP-1999; 93DK-0001340.
 XX XX
 PA (NOVO) NOVO NORDISK AS.
 XX XX
 PI Lehmann M;
 XX XX
 DR WPI: 2000-491161/43.
 XX N-PSDB; AAA73231.
 XX XX
 PT Novel phytases with improved properties such as temperature stability,

PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX Example 9; Fig 2a-c; 240pp; English.
 XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

Sequence 467 AA;

Query Match 93.6%; Score 2311; DB 21; Length 467;
 Best Local Similarity 93.1%; Pred. No. 3.8e-224;
 Matches 435; Conservative 13; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGATLPGRNHSCDVTGQYQCFPEISHLWGQYSPFFSLADE 60
 DB 1 mgvfvlisiatlfgstgatalprgnshscdvtggyqcfpeishlwgqyspyfslade 60
 QY 61 SAISPDVPKGRVTFVQLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKTYN 120
 DB 61 saispdvpdkgrvtfvqlsrhgaryptsskksaysalieaigknatafkgyafikty 120
 QY 121 YTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAEKFTGQFSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalkarkivpfirasgsdrviasaekfiegfsa 180
 QY 181 KLADPGANPHQASPVINVIIEPEGAGYNTLDHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
 DB 181 kladpganphqaspvdiviiepegsgyntldhgtctafedselgddveanftalfapair 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDCPEDTAVRTSDATOLSPFCDLTHDEWTOYDYLSLG 300
 DB 241 arleahlpgvnltdedvvylnmdcpfetavrtsdatslpcalfthdewrqdyylsg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMYSIFALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARYVEMMOCEAKEP 420
 DB 361 hdnsmisifalglngtaptlsttsvesieetdgyaswtvpfgarayvemmqcqaek 420
 QY 421 LVRVLVNDVRVPLHGGVCKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlvndrvplhgcavdklgrckrddfveglsfarsggnwaecfa 467

Search completed: October 26, 2001, 16:40:03
 Job time: 4954 sec

us-09-488-265-26.rag

Sat Oct 27 15:25:38 2001

...

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: October 26, 2001, 14:11:09 ; Search time 75.85 Seconds
(without alignments)
814.587 Million cell updates/sec
Title: US-09-488-265-26
Perfect score: 2470
Sequence: 1 MGVFVLLSIATLFGSTSGT.....DFVEGLSFARSGNWECEFA 467
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1902	77.0	465	3	O00092	O00092 aspergillus
2	1861	75.3	467	3	O9UJZ7	O9UJZ7 aspergillus
3	1855	75.1	467	3	O93838	O93838 aspergillus
4	1842	74.6	467	3	O9HEQ0	O9HEQ0 aspergillus
5	1826	73.9	466	3	O00100	O00100 aspergillus
6	1756	71.1	466	3	O00085	O00085 aspergillus
7	1719	69.6	466	3	O00096	O00096 talaromyces
8	1346.5	54.5	487	3	O00107	O00107 thielavia h
9	343.5	13.9	469	3	O9Y846	O9Y846 kluyveromyc
10	333.5	13.5	442	3	O74677	O74677 pichia anqu
11	301	12.2	463	3	O60172	O60172 schizosacch
12	212.5	8.6	460	5	O9VV72	O9VV72 drosophila
13	204	8.3	451	11	O35217	O35217 rattus norv
14	204	8.3	467	5	O96421	O96421 drosophila
15	196	7.9	481	11	O92216	O92216 mus musculu
16	195	7.9	487	4	O95172	O95172 homo sapien
17	194	7.9	487	4	O9UNW1	O9UNW1 homo sapien
18	192	7.8	487	4	O9UGA3	O9UGA3 homo sapien
19	179	7.2	449	13	O92170	O92170 gallus gall

20	176.5	7.1	453	5	O96420	O96420 drosophila
21	175.5	7.1	453	5	O9W438	O9W438 drosophila
22	167.5	6.8	198	3	O9UTX1	O9UTX1 schizosacch
23	154.5	6.3	468	10	O04509	O04509 arabidopsis
24	144	5.8	274	11	O9JUD5	O9JUD5 mus musculu
25	130	5.3	449	5	O19076	O19076 caenorhabdi
26	129.5	5.2	374	11	O9JMG5	O9JMG5 mus musculu
27	128.5	5.2	381	11	O9QXG5	O9QXG5 mus musculu
28	119.5	4.8	513	2	O46334	O46334 comamonas t
29	119	4.8	380	5	O22525	O22525 caenorhabdi
30	111.5	4.5	827	2	O24781	O24781 eubacterium
31	110.5	4.5	381	11	O9QXH7	O9QXH7 mus musculu
32	109	4.4	683	5	O00838	O00838 leishmania
33	109	4.4	707	5	O00839	O00839 leishmania
34	108.5	4.4	513	2	O24719	O24719 comamonas t
35	108	4.4	354	5	O19709	O19709 caenorhabdi
36	106.5	4.3	730	5	O20826	O20826 caenorhabdi
37	104	4.2	499	3	P87212	P87212 polyporacea
38	103.5	4.2	496	1	O34184	O34184 halobacteri
39	102.5	4.1	1013	14	O9ILX9	O9ILX9 retroperito
40	101.5	4.1	447	5	O9U5U0	O9U5U0 drosophila
41	101.5	4.1	447	5	O9U5T8	O9U5T8 drosophila
42	101.5	4.1	447	5	O9TW19	O9TW19 drosophila
43	101.5	4.1	602	2	O9K5A0	O9K5A0 erwinia chr
44	101	4.1	516	5	O25327	O25327 leishmania
45	101	4.1	763	2	O9R6Y7	O9R6Y7 anabaena sp

ALIGNMENTS

RESULT 1
O00092
ID O00092 PRELIMINARY; PRT; 465 AA.
AC O00092;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHVA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN=ATCC 34625;
RX MEDLINE=97288063; PubMed=9143104;
RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
RT phytase from the fungus Aspergillus fumigatus.";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMATIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC EMBL; U59804; AAB96872.1; -
CC HSSP; P34752; 1IHP.
CC InterPro; IPR000560; -
DR Pfam; PF00328; acid.phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 26

```

FT CHAIN 27 465 3-PHYTASE A.
FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
FT SIMILARITY).
FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 359 359 PROTON DONOR (BY SIMILARITY).
FT DISULFID 30 39 BY SIMILARITY.
FT DISULFID 70 412 BY SIMILARITY.
FT DISULFID 213 463 BY SIMILARITY.
FT DISULFID 262 280 BY SIMILARITY.
FT DISULFID 434 442 BY SIMILARITY.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 50836 MW; 86FCID9058C9B2C9 CRC64;

```

```

Query Match 77.0%; Score 1902; DB 3; Length 465;
Best Local Similarity 77.1%; Pred. No. 2.5e-147;
Matches 360; Conservative 37; Mismatches 68; Indels 2; Gaps 2;

```

```

QY 1 MGVEVLLSLATLFGSTGTLGPRGNHSCDTVDGGYOCFPEISHLWGOYSPFFSLADE 60
DB 1 MVTLTLLSAVLLSGRVSAAPSSAG-SKSCDTVDLGYQCSPTATSHLWGOYSPFFSLEDE 59
QY 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFLKTYN 120
DB 60 LSVSSKLKDCRITLQVLSRHGARYPTSSKSKYKLVTAIQANATDFKGFALKTYN 119
QY 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
DB 120 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 179
QY 181 KLADPGANPHQASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 180 KLADPGAT-NRAAPASVVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 238
QY 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSDFCDLTHDEWIOYDYLQSLG 300
DB 239 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSDFCDLTHDEWIOYDYLQSLG 298
QY 301 KYGYGAGNPLGPAQGVGFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
DB 299 KYGYGAGNPLGPAQGVGFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 358
QY 361 HDNTMVSIFPFGALGYNKTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAKEP 420
DB 359 HDNTMVSIFPFGALGYNKTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAKEP 418
QY 421 LVRVLVNDRVVPLHGGCVGDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
DB 419 LVRVLVNDRVVPLHGGCVGDKLGRCKRDDFVEGLSFARSGGNWEECEFA 465

```

```

RESULT 2
Q9U0Z7 PRELIMINARY; PRT; 467 AA.
ID AC Q9U0Z7
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 15, Last annotation update)
DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLIASE PRECURSOR (EC 3.1.3.8).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RA Hongning W., Qi W., Jing X.;
RT "PCR, cloning and characterization of the phytase (phyA) gene of

```

```

RT Aspergillus niger (China Strain).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218813; AAF25481.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Signal; Lyase; Hydrolase.
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

```

```

Query Match 75.3%; Score 1861; DB 3; Length 467;
Best Local Similarity 74.3%; Pred. No. 5.6e-144;
Matches 347; Conservative 47; Mismatches 73; Indels 0; Gaps 0;

```

```

QY 1 MGVEVLLSLATLFGSTGTLGPRGNHSCDTVDGGYOCFPEISHLWGOYSPFFSLADE 60
DB 1 MGVSALLPLLYLLSGVTSGGLAVPASRNQSTCDTVDOGYOCFSETSHLWGOYAPFFSLANE 60
QY 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFLKTYN 120
DB 61 SATSPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNATAFKGYAFLKTYN 120
QY 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
DB 121 YTLGADDLTPFGGQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIAAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIIEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
DB 181 KLADPGAPQGSPPKIDVISEASSNNNTLDPGTCTVFDESELADAVEANFTAVFPTIR 240
QY 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSDFCDLTHDEWIOYDYLQSLG 300
DB 241 ARLEAHLPGVNLDEVDVNLMDCPDFTVARTSDATQLSDFCDLTHDEWIOYDYLQSLG 300
QY 301 KYGYGAGNPLGPAQGVGFVNEIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
DB 301 KYGYGAGNPLGPTGGVGYANIELIARLTHSPVHDDTSSNHTLDSNSATFPLNSTLYADFS 360
QY 361 HDNTMVSIFPFGALGYNKTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAKEP 420
DB 361 HDNGLIISLTFALGYNKTKPLSTTVQNTTQDTGFSANTVFPASRLVEMMOCEAKEP 420
QY 421 LVRVLVNDRVVPLHGGCVGDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
DB 421 LVRVLVNDRVVPLHGGCVGDKLGRCKRDDFVEGLSFARSGGNWEECEFA 467

```

```

RESULT 3
O93838 PRELIMINARY; PRT; 467 AA.
ID AC O93838
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PHYTASE.
GN PHVA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-57;
RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
RT "Phytase having high-affinity for phytic acid";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022700; CAB19824.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.

```

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543EDC265C CRC64;

Query Match 75.1%; Score 1855; DB 3; Length 467;
 Best Local Similarity 73.4%; Pred. No. 1.7e-143;
 Matches 343; Conservative 51; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGCTALPGRNHSCDVTDDGGYQCPEISHLWGOYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNOSTCDTVDGGYQCFSETSHLWGOYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFAQVLSRHGARYPTDSKKYSALIEIEIQNATTFEKGYAFKLTYN 120
 QY 121 YTLGADDLTPFGEQMVNSGKIFRYRYKALARKIVPVFRASGSDRVIASAEKFIQGFOSA 180
 DB 121 YSLGADDLTPFGEQELVNSGVKIFYRVESTRNIVPFISSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGANPHOASPVINIIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGOSSPKIDWVISEASTSNNTLDPGCTVFEDSELADDEANFTATVPSPIS 240
 QY 241 ARLEAHLPGVNLTDVNVNMDMCPFDVARTSDATQLSPFCDLTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVSLTDEVTYVLDNMCSEDTISTSTVDTKLSPFCDLTHDEWINYDYLQSLN 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPTQGVGYANELLARLTHSPVHDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HDNTMVSIFGALYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISLTFALYNGTKPLSSTTAENITQDGFSSARTVPFAARMYVEMMOQCEQEP 420
 QY 421 LVRVLNDRVPLHCGGVDKLGRCKRDDFVEGLSFARSGGNWEECFA 467
 DB 421 LVRVLNDRVPLHCGPVDALGRCTRDSEFKGLSFARSGGDWGECEFA 467

RESULT 4

ID Q9HEQ0 PRELIMINARY; PRT; 467 AA.
 AC Q9HEQ0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE PHYTASE.
 OS Aspergillus ficuum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L., An L., Wang Y., Yuan X.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY013315; AAG40885.1; -;
 SQ SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match 74.6%; Score 1842; DB 3; Length 467;
 Best Local Similarity 73.2%; Pred. No. 2e-142;
 Matches 342; Conservative 51; Mismatches 74; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGCTALPGRNHSCDVTDDGGYQCPEISHLWGOYSPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSTCDTVDGGYQCFSETSHLWGOYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKKYSALIEAIQKNATAFKGYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFAQVLSRHGARYPTDSKKYSALIEIEIQNATTFEKGYAFKLTYN 120

QY 121 YTLGADDLTPFGEQMVNSGKIFRYRYKALARKIVPVFRASGSDRVIASAEKFIQGFOSA 180
 DB 121 YSLGADDLTPFGEQELVNSGVKIFYRVESTRNIVPFISSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGANPHOASPVINIIPEGAGYNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 KLKDPRAQPGOSSPKIDWVISEASTSNNTLDPGCTVFEDSELADDEANFTATVPSPIS 240
 QY 241 ARLEAHLPGVNLTDVNVNMDMCPFDVARTSDATQLSPFCDLTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVSLTDEVTYVLDNMCSEDTISTSTVDTKLSPFCDLTHDEWINYDYLQSLN 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPTQGVGYANELLARLTHSPVHDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HDNTMVSIFGALYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP 420
 DB 361 HDNGIISLTFALYNGTKPLSSTTAENITQDGFSSARTVPFAARMYVEMMOQCEQEP 420
 QY 421 LVRVLNDRVPLHCGGVDKLGRCKRDDFVEGLSFARSGGNWEECFA 467
 DB 421 LVRVLNDRVPLHCGPVDALGRCTRDSEFKGLSFARSGGDWGECEFA 467

RESULT 5

ID 000100 PRELIMINARY; PRT; 466 AA.
 AC 000100;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE 3-PHYTASE PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE).
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 116.46;
 RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B.,
 RA Broger C., van Loon A.P.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U60412; AAB58465.1; -;
 DR HSSP; P34752; IHP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 466 3-PHYTASE.
 FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY
 FT SIMILARITY).
 FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 31 40 BY SIMILARITY.
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	207	207	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	376	376	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SQ	SEQUENCE	466 AA;	51055 MW;	F2AECBCIAF7C22C4	CRC64;
Query Match					
Best Local Similarity 73.9%; Score 1826; DB 3; Length 466;					
Matches 339; Conservative 44; Mismatches 83; Indels 0; Gaps 0;					
QY	1	MGVFWLLSIATLFGSTGTALGPRGNSHSDTVGGYQCFPEISHLWGOYSPFSLADE	60		
DB	1	MGVFWLLSIATLFGSTGTALGPRGNSHSDTVGGYQCFPEISHLWGOYSPFSLADE	60		
QY	61	SAISPDVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFLKTYN	120		
DB	61	SPFPLDVPDCHITFVQVLSRHGARYPTSSKSKYSALIEATQKNATAFKGYAFLKTYN	120		
QY	121	YTLGADDLTPFGEQOMVNSGIGKIFRYRKALARKIVPFVRASGDRVIAAEKFEQFSA	180		
DB	121	YSGSENLNPFGRNQLDLGAOFYERYNALTFRHINPFRATDASRVHESAEKFEQFSA	180		
QY	181	KLADPGANPHQASPVNIIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP	240		
DB	181	ROGDHANPHQSPRVDDVVIPEGTAYNNITLHSHCTAFESTVGDAAADFTAVFAP	240		
QY	241	ARLEAHLPGVNLTDDEVNLMDCPFTVARTSDATQSLSPFCDLFTHDEWIQDYLSLG	300		
DB	241	KRLAADLPGVNLTDDEVNLMDCPFTVARTSDATQSLSPFCDLFTHDEWIQDYLSLG	300		
QY	301	KYGYGAGNPLGPAQGVFVNLIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	360		
DB	301	KYGYGAGNPLGPAQGVFVNLIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADF	360		
QY	361	HNTWVSIFWALGLYNGTAPLSQTSVESVSDTGDYAAATVFPFAARAYVENMOCRA	420		
DB	361	HNTWVSIFWALGLYNGTAPLSQTSVESVSDTGDYAAATVFPFAARAYVENMOCRA	420		
QY	421	LVRVLNDRVVPVPHGCGVDKLGCRKRDDEVEGLSFARSGGNWEECF	466		
DB	421	LVRVLNDRVVPVPHGCGVDKLGCRKRDDEVEGLSFARSGGNWEECF	466		
RESULT 7					
000085					
ID	000085	PRELIMINARY;	PRT;	466 AA.	
AC	000085				
DT	01-JUL-1997	(T-EMBLrel. 04, Created)			
DT	01-JUL-1997	(T-EMBLrel. 04, Last sequence update)			
DT	01-MAY-2000	(T-EMBLrel. 13, Last annotation update)			
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-				
DE	PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE				
GN	PHYA.				
OS	Aspergillus terreus.				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;				
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.				
OX	NCBI_TaxID=33178;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=9A1;				
RX	MEDLINE=97117792; PubMed=9025298;				
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;				
RT	"the phytase subfamily of histidine acid phosphatases: isolation of				
RT	genes for two novel phytases from the fungi Aspergillus terreus and				
RT	Myceliophthora thermophila.";				
RL	Microbiology 143:245-252(1997).				
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE				
CC	FROM PHYTATE (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-				
CC	INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.				

000096;
 01-JUL-1997 (TReMBLrel. 04, Created)
 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 PHSA.
 GN PHSA.
 OS Talaromyces thermophilus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Talaromyces.
 OX NCBI_TaxID=28565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 20186;
 RX MEDLINE=98007872; PubMed=9349716;
 RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,
 RA van Loon A.P.G.M.;
 RT "Cloning of the phytases from *Emericella nidulans* and the thermophilic
 fungus *Talaromyces thermophilus*.";
 RL Biochim. Biophys. Acta 1353:217-223(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- MISCELLANEOUS: HAS POTENTIAL AS FOOD AND FEED ADDITIVE. IT CAN
 FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
 USED AS FOOD FOR MONOGASTRIC ANIMALS.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59802; AAB96873.1; -
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; -
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 DR Hydrolase; Glycoprotein; Signal
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 1 466 3-PHYTASE A.
 FT ACT_SITE 79 79 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 357 357 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 28 37 BY SIMILARITY.
 FT DISULFID 68 410 BY SIMILARITY.
 FT DISULFID 212 461 BY SIMILARITY.
 FT DISULFID 261 278 BY SIMILARITY.
 FT DISULFID 432 440 BY SIMILARITY.
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51450 MW; FC4575B521A5C929 CRC64;
 Query Match 69.6%; Score 1719; DB 3; Length 466;
 Best Local Similarity 69.3%; Pred. No. 2.3e-132;
 Matches 325; Conservative 45; Mismatches 91; Indels 8; Gaps 3;
 QY 1 MGFWVLLS--IATLFGSTGALPGRNSHSDTVGGYQCPPEISHLWGQSPFSLA 58
 DB 1 MSLLLVLSGLVALYVSRN----PHVDSHSCNTVEGGYQCPPEISHWGQSPFSLA 55
 QY 59 DESAISPDPKGRVTFVQVLVSRHGRYPSTSSKYSALIEAIQKNATAPKGYAFLEKT 118
 DB 56 DQSEISDPVQNKITFVQLVSRHGRYPSTSSKYSALIEAIQKNATAPKGYAFLEKT 115
 QY 119 YNTLGAADLLTPGEQMVNSGKIFRYRYKALARKIVPFVRASGSDRVIASAKFIEGFQ 178
 DB 116 YRQLGANDLTPGENQMIQGLKFNHYKSLARNAPFVRCSGSDRVIASGRLFIEGFQ 175
 QY 179 SAKLADPGANPHQASPVINVIIEGAGYNNLTDLHGCLTAPEESELGDDVEANFTAVFAPP 238
 DB 176 SAKVLDPHSKDHPPTINVIIEEGSYNNLTGSCPFVEDSSGGHDAOEKAKQFAPA 235

QY 239 IRARLEAHLPGVNLTDQVNVNMDMCPEDTVARTSDATQLSFPDCLFTHDEWIQDYDLOS 298
 DB 236 ILEKIKDLPGVDLAVSDVPYLMDCLPETLARNHTDT-LSPFCALSTQEEWQAYDYQS 294
 QY 299 LKYYGYGAGNPLGPAQGVGFVNELIARTLTHSPVDHTSTNHTLDSNPNATFPLNATLYAD 358
 DB 295 LKYYGNGGNGPLGPAQGVGFVNELIARTLTHSPVDHTSTNHTLDSNPNATFPLNATLYAD 354
 QY 359 FSHDNTMVSITFALGLYNGTKPLSTSVESIEETDGYAASWTVPFAARAYVEMMOCEAK 418
 DB 355 FSHDNTMVSITFALGLYNGTKPLSTSVESIEETDGYAASWTVPFAARAYVEMMOCEAK 414
 QY 419 EPLVRLVNDVRVPLHGGVCKLGRCKRDDFVEGLSFARSGGNWEECEFA 467
 DB 415 EPLVRLVNDVRVPLHGGVCKLGRCKRDDFVEGLSFARSGGNWEECEFA 463
 RESULT 8
 ID 000107 PRELIMINARY; PRT; 487 AA.
 AC 000107;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 PHSA.
 GN PHSA.
 OS Thielavia heterothallica.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Chaetomiaceae; Thielavia.
 OX NCBI_TaxID=78579;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97177792; PubMed=9025298;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,
 RA van Loon A.P.G.M.;
 RT "The phytase subfamily of histidine acid phosphatases: isolation of
 genes for two novel phytases from the fungi *Aspergillus terreus* and
Myceliophthora thermophila.";
 RL Microbiology 143:245-252(1997).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO
 ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA
 SHIFTED TO MORE ACIDIC PH VALUES.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59806; AAB52508.1; -
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; -
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 75 POTENTIAL.
 FT CHAIN 1 487 3-PHYTASE A.
 FT DOMAIN 267 270 POLY-SER.
 FT DOMAIN 423 433 POLY-GLY.
 FT ACT_SITE 75 75 REQUIRED FOR BINDING SUBSTRATE (BY
 SIMILARITY).
 FT ACT_SITE 76 76 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 368 368 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 26 35 BY SIMILARITY.
 FT DISULFID 64 421 BY SIMILARITY.
 FT DISULFID 208 485 BY SIMILARITY.
 FT DISULFID 260 289 BY SIMILARITY.
 FT DISULFID 456 464 BY SIMILARITY.
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	CARBOHYD	247	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	346	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	487 AA; 52537 MW; 97D10EDC83D051DB CRC64;	

Query Match	54.5%;	Score 1346.5;	DB 3;	Length 487;
Best Local Similarity	53.9%;	Pred. No. 7.3e-102;		
Matches 268;	Conservative	58;	Mismatches 128;	Indels 43;
Gaps				
Qy	1	MGVFFVL---LSTATLFGSTSGTALGPRGNHSCDTPVDGQCFPIEISHLWGQYSPFSL	57	
Db	4	LGWVMVWGFLATASL-----QSESRPCDTPDLGFCQGTALSHFWGQYSPFSV	52	
Qy	58	ADESAISPDVPKGCRTFFVOVLSRHGARPYPTSSKKYSALIEAQKNATAFKGYAFIL	117	
Db	53	PSE--LDASIPDDCEVTFAQVLSRHGARAPTLKRAASYVDLIRHHGAISYCPGYEFLR	110	
Qy	118	TYNYTLGADDLTFEGEOOMVNSGIKFYRRYKALARKIVFPVRASGSDRVIASAEKIEGF	177	
Db	111	TYDYLGADELTTGQOOMVNSGIKFYRRYRALKARKSIPFVRTAGQDRVVHSAENTOGF	170	
Qy	178	QSAKLADPGANPHQASPVINVIIPGAGYVNTLDHGLCTAFEE---SELGDDVEANFTAV	234	
Db	171	HSALLADRGSTVRPTLYDMWVIPETAGANTLHNDLCTAFEEGPSTIGDDAQDTYLS	230	
Qy	235	FAPPIRARLEAHLPGVNLITDEDVYNLMDCMPFDFTVARTS-----DAFQLSPFCD	283	
Db	231	FAGPIITARVNANLPGANLTDADTVALMDLCPETVAVSSSDPATADAGGNGRPLSPF	290	
Qy	284	LFTHDEWIDYDYLQSLGKYYGYGAGNPLGPAQGVGFVNELIARLTHISPVQDHTSTNHTLD	343	
Db	291	LFSESEWRAYDYLQSGVKWYGYGPNLPGTQGVGFVNELIARLTHISPVQDHTSTNHTLD	350	
Qy	344	SNPATFPLNATLYADFSDHNTWYSIFPALGLXNGTKPLSTTSVESTETDGYAASMTVPF	403	
Db	351	GDPRTFPLGRPLYADFSDHNDMMGVLGALGAYGVPPDLKTARDDPEELGGYAAASNAVPE	410	
Qy	404	AARAYVEMMOC-----EAEKEPLRVRLVNDVRVPLHGGCVKDLGKRCKRDEVF	450	
Db	411	AARIYVEKMRCSGGGGGGGEGQEKDEEWRVRLVNDVRVMTLKGCGADERGMCITLERFI	470	
Qy	451	EGLSFARSGGNWEECFA	467	
Db	471	FSMAFARGNGKWDLCFA	487	

RESULT	9	
Q9Y846		
ID	PRELIMINARY;	PRT; 469 AA.
AC	Q9Y846;	
DT	01-NOV-1999 (TREMblrel. 12, Created)	
DT	01-NOV-1999 (TREMblrel. 12, Last sequence update)	
DT	01-MAR-2001 (TREMblrel. 16, Last annotation update)	
DE	ACID PHOSPHATASE (EC 3.1.3.2).	
GN	PHO3.	
OS	Kluyveromyces lactis (Yeast).	
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
OC	Saccharomycetales; Saccharomycetaceae; Kluyveromyces.	
OX	NCBI_TaxID=28985;	
[1]	SEQUENCE FROM N.A.	
RP	STRAIN=2360/7;	
RA	San Vicente A., Ferminan E., Dominguez A.;	
RT	"Isolation and characterization of KlPHO3 a gene encoding a	
RT	constitutive acid phosphatase from Kluyveromyces lactis.";	
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AJ007502; CAB46490.1;	
DR	HSSP; P34755; 10FX.	
DR	InterPro; IP000560;	
DR	Pfam; PF00328; acid_phosphat; 1.	
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.	
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.	
KW	Hydrolase.	

FT CHAIN 17 469 ACID PHOSPHATASE.
SQ SEQUENCE 469 AA: 52560 MW: 5C7ABF622CEA891C CRC64;

Query Match 13.9% Score 343.5 DB 3 Length 469;

Best Local Similarity	25.18	Pred. No. 7.7e-20;	
Matches	123; Conservative	68; Mismatches	209; Indels
			91; Gaps
			17;

```

Qy 7 LLSTATLFGSTGALTGPRGNHSHCDTVDGGYOCF-----PETSHLWGQYSPPFSLA 58
Db 1 MLSILLLSLSUGTHAAPISK-----DMGTVCYALNSTTDESFPLLNGOGPHYDYP 53
Qy 59 DESAISDPVKGRCVTFVQVLSRHGARYPTSSKSKYSALIEAIOKNAFAFKGYAFLKT 118
Db 54 QSFQIPVEVPDQCTVEHVQMLARHGERTASKGLKWLIALWDKLEKFOGQYNGPMEVFN 113
Qy 119 YNY-----TLGADDLT-----PP-GEQOMVNSGTFKFRYKALARKIVPVRASGSD 164
Db 114 YEFTVSNTKYFDQUTNSTVDVDPNYPAGATAOHLGKIYAYNGDLFSQNP-VFTSSSG 172
Qy 165 RVIASAKFIEGFSAKLADPGANPHOASPVIN-----IPEG--AGYNNTLDHGLCTA 217
Db 173 RVHQTAKVYVSSLEE-----LDLDLQIIQENETSGANSLTPADSCMT 217
Qy 218 FEESLGDDVEANFTAVFPAPPIRARLEAHPGVNLT--DEDVNLMDMCPFDFTVARTSDA 275
Db 218 Y-NGDLGDEVFENATLPYLTDIKNRWKKNSNLNLTLEHDDIELLDWCACAFETNVAGSSA 276
Qy 276 TQLSPFCDLFTHDEWIQVDYILQSLGKYGYGACGNPLGPAQGVGFVANELIARLTHSPVQDH 335
Db 277 V-----CDLFRNDLVAYSYVYANNVFNRRGAGNPMSPNIGSVLVNVAAYNLTLQADBLDN 331
Qy 336 TSTNHTLDSNPATFPLNATLYADESHDNTWVSFFALGLY-NGTKPLSTTTSVESIETDG 394
Db 332 -----KWLSESHDTDIOQFISALGLIDNG---VTEYSLDQOVDFOQNI 370
Qy 395 YAASWTVPFAARVYVEMMOCEAEPLVRLVNDRVVPLHGCVGDKLGRCKRDDF----- 449
Db 371 QQLSWVTMPGRIETKLC-----GNASVRYIYINDVIIPVPGCISGPGFCIEDFDYIT 438
Qy 450 --VEGLSFARS 458
Db 429 NRLNGIDYVSS 439

```

RESULT	10	
074677		
ID	PRELIMINARY;	PRT; 442 AA.
AC	074677;	
DT	01-NOV-1998 (TrEMBLrel. 08, Created)	
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)	
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)	
DE	REPRESSIBLE ACID PHOSPHATASE (EC 3.1.3.2).	
GN	PHO1.	
OS	Pichia angusta (Yeast) (Hansenula polymorpha).	
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	
OC	Saccharomycetales; Saccharomycetaceae; Pichia.	
OX	NCBI_TaxID=4905;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-ATCC34438;	
RX	MEDLINE=98386672; PubMed=9720203;	
RA	Phongdara A., Merckelbach A., Keup P., Gellissen G., Hollenberg C.P.,	
RT	"Cloning and characterization of the gene encoding a repressible acid	
RT	phosphatase (PHO1) from the methylotrophic yeast Hansenula	
RT	polymorpha."	
RL	Appl. Microbiol. Biotechnol. 50:77-84(1998).	
DR	EMBL; AF051161; AAC62537.1; -	
DR	HSSP: P34755; 10FX.	
DR	InterPro: IPR000560; -	
DR	Pfam: PF00328; acid_phosphat; 1.	
DR	PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.	
KW	Hydrolase.	

```
SQ SEQUENCE 442 AA; 49370 MW; 7087D91A85B05C31 CRC64;

Query Match 13.5%; Score 333.5; DB 3; Length 442;
Best Local Similarity 25.8%; Pred. No. 4.6e-19;
Matches 106; Conservative 63; Mismatches 159; Indels 83; Gaps 16;

QY 66 DVPKGRVTFVQVLSRHGARYPTSSKSKYSALIEAIQK-NATAFKGKYAFLLKTYNTLIG 124
DB 55 DTPPHCEIEQAQLMRHGERPTKSSGGQKKFYDKLAKANITDYKGLPIAFIEDLEYFVP 114
QY 125 ADDLTPFGEQWVNSGI-----KF-----YRRYKAL--ARKIVFVRASGSDRVIASAEKFI 174
DB 115 DSDNYELETTRGLYSLGLNAPKFGTYRLERYDSLVDTSVLPIFAAS-EDRVVDIARSGF 173
QY 175 EGFQSAKLADPGANPHQASPVINVI-----IPEGAGYNTLDHGLC-----TAFESSELGDD 226
DB 174 RGFPGPDYA-----TSCSIQVNETDTSGKANALATKDN--CPTVNSFFYDYSFGDE 223
QY 227 V---EANTAVFAPPARLEAHLPGVNLTDVNVNMDMCPFDTVARTSDATQLSPFCD 283
DB 224 IFQREAD-----RLNELSPGFNITADDITMGTYCAYETNVKGH-----SSPCD 267
QY 284 LFTDHEWIOYDLSGLGKYCYGAGNPLGPAQGVGVNELLIALTLHSPVQDHTSTNHTLD 343
DB 268 ALSREAFALQYNDVTKFYQFGCYNMSAVAGGYAN----- 305
QY 344 SNPAPFPL--NATLYADFSDHNTMWSIFFALGNGYTKPLSTTSVESIETDGYAASWT 400
DB 306 ---ATAKLQEDGKLWFSFSDHNDLLNVIYALGLITDFE-----LGTEDVDFHRFKTSEL 358
QY 401 VPFARAYVEMMOCEAKEPLRVVLVNDRVVPLHCGVDKLGRCRDRDFVE 451
DB 359 VPOGARLIIKLN--SDTSEVRTILNDKVVYPVPGCSGGPGYSCPLEDYLD 407

RESULT 11
060172 ID O60172 PRELIMINARY; PRT; 463 AA.
AC O60172;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE PROBABLE ACID PHOSPHATASE (EC 3.1.3.2).
GN SPC21H7.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; AL023286; CAAL18863.1; -.
DR HSSP; P34755; 1QFX.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein.
FT ACT_SITE 69 69 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 340 340 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```
SQ SEQUENCE 463 AA; 52758 MW; 6C41AE422C6D624A CRC64;

Query Match 12.2%; Score 301; DB 3; Length 463;
Best Local Similarity 24.3%; Pred. No. 2.3e-16;
Matches 106; Conservative 71; Mismatches 198; Indels 52; Gaps 13;

QY 51 YSPFSLADESAISDPVPGCRVTFVQVLSRHGARYPTS--SKSKKYS-----L 98
DB 43 HEPYFDGLDSA-----FPETCEIQOQHLLQRHGSRNPTGDVTATDVYSOYLNNFOEKL 97
QY 99 IEAIQKNATAFKGKYAFLLKTYNTL--CADDLTPFGEQWVNSGIKFYRRYKALKARKIV 155
DB 98 NGSPVNSYFENPCFKQWTPVIDAENADOLSRGRLEUFDLGRQLRYRYKLFDSYV 157
QY 156 PFVRASGSDRVIASAEKFIETGFSAKLADPGANPHQASPVINVIPEG--AGYNTLTDHG 213
DB 158 YDINTAEQERVVESAKWFTYGLFGDKMYE-----KTNFILLSEKGAAGANSLSMYN 208
QY 214 LCTAFEESEL-----GDDVEANFTAVFAPPARLEAHL--PGVNLTDVNVNMDMCPFD 268
DB 209 ACPVFKDNFNHKNATDAAHAVVRNFIIEPIVNLAKYFDSSYKLTINDVRSIFYICEVEI 268
QY 269 VARTSDATQLSPFCDLFTDHEWIOYDLSGLGKYCYGAGNPLGPAQGVGVNELLIALRLT 328
DB 269 AIKDH-----SDFCSIFTPSEFLNFEYDSDLDQAYGGFVSEWASTLGGAYINNLA 322
QY 329 HSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMWSIFFALGNGYTKPLSTTSVE 388
DB 323 -----RNVTPDPEDRK-----VFLATHDSNIIPVEAALGFFPDITPQNPLPTDK 367
QY 389 IEEWDGVAASWTVPFAARAYVEMMOCEAKEPLRVVLVNDRVVPLHCGVDKLGKGR---- 444
DB 368 NIYTSOKTSSSFVPFAGNLITELFFC--SDSKYYVRHLVNOQVYPLIDCGYSGTSGDLC 426
QY 445 KRDDFVEGLSELRGGN 461
DB 427 ELQAYLNSPIRANSTSN 443

RESULT 12
Q9VVW72 ID Q9VV72 PRELIMINARY; PRT; 460 AA.
AC Q9VV72;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
DE MIPPI PROTEIN.
GN MIPPI OR CG4123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabros B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
```


Db 131 DLFCRENYGRRLITSSKRCVDSAAFLQGLW--QVHHPGLPP---PDVSDMECDPPR 184
 QY 206 YNNITL---DHGLCTAFESSELDGDDVEANFTAVF-----APPRARLE-----AHLPG 249
 Db 185 VNDKLMRFEDH--CKFL-----TEVERNATLYHVEAFKTPGEMQTLKKVAATLQVPV 237
 QY 250 VNLTDEDVNNLMDGPFDTVARTSDATOLSPCDLFTHDEWIQYDYLQSLGKYGYGAGN 309
 Db 238 NNLNADLIQVAFCTGSDLAIOGVH---SPWCDVFDVDDAKVLEYLNDLKOYWRKSYGY 293
 QY 310 PLGPAQGVGVFNELLARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIF 369
 Db 294 AINSRSCNLFQDFLHLDKAVEQKORSQ-----PVSSVILQFQHAETLLPLL 342
 QY 370 FALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAEKEP-----LVRVL 425
 Db 343 SLMGYFKDKPLTAYNFEQVHRE-FRSGHIVPYASNLIFVLYHCEDAQTPOEKFQIOML 401
 QY 426 VNRDVVPL 433
 Db 402 LNEKVLP 409

RESULT 14
 ID O96421 PRELIMINARY; PRT; 467 AA.
 AC O96421;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 1.
 GN MIPPI OR CG4123.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046913; AAD02436.1;
 DR FlyBase; FBgn0026061; Mippl.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00034; ER_TARGET; UNKNOWN_1.
 SQ SEQUENCE 467 AA; 53602 MW; 1A0571B9958B2632 CRC64;

Query Match 8.3%; Score 204; DB 5; Length 467;
 Best Local Similarity 22.9%; Pred. No. 2e-08;
 Matches 103; Conservative 66; Mismatches 173; Indels 108; Gaps 20;

QY 66 DVPK-----GCRVTFVQVLSRHGARYPTSS---KSKYSALIEAIQKNATAFKKGK-----112
 Db 46 DIDKQVLPVPCQPKMWIFRHGTRLPKPKSMINKASRVAELRDLIINNQQVARTKPTDA 105
 QY 113 -----YAEFLKTYNTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVPV-----158
 Db 106 LCQTDLIAIKLWKN---SSITPDMEEYLTAGQYEDLRGTAKLYQRYTTLTANTNDY 162
 QY 159 ---RAGSDRVIASAEKFTG-FQSAKLADPGANPHQ-----ASPVINY-IIPEGA 204
 Db 163 YQFRHTDTORTTESKFAEGLFGSQAHPVEIPKQDLLLRPYDYCSSFKVNYKDEGS 222
 QY 205 GYNNITLDHGLCTAFESSELDGDDVEANFTAVFAPPRARLEAHLPG--VNLTDEDVNNLMD 262
 Db 223 EYK-----FHQSCLYNDTLAD-----ISTRL-----GFLYTLERADIKLYMD 260
 QY 263 MCPFDVARTSDATOLSPCDLFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVFN 322
 Db 261 MCRYE---QAMVNRDNRVWCGAFLEQITVFYEYLDLKYGYGFFPENALNCRVLQD 317

QY 323 LIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKPLS 382
 Db 318 LLTHLS-NPVSPH-----VXHFGHSTGLLLTLLTALGIXKDDIKUR 357
 QY 383 TTSVESIETDGYAASWTVPFAAR--AYVEMMOCEAEKEPLVRLVNDRVVPLHGGGVDK 440
 Db 358 ADNYDSL-TSRWKSLLDPFAANEVAVKYDLFADLDREKVV-FELNQOAVQLDWCSV--413
 QY 441 LGRCKRDRDFVE-----GLSFARSGG 460
 Db 414 -GLCKWSDVLEKYKTIADACGCEYCYRTGG 442

RESULT 15
 ID O922L6 PRELIMINARY; PRT; 481 AA.
 AC O922L6;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
 GN MINPPI OR MIPP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046908; AAD02434.1;
 DR MGD; MGI:1336159; Minppl.
 DR InterPro; IPR000560; -
 DR InterPro; IPR000886; -
 DR Pfam; PF00328; acid_phosphat; 2.
 DR PROSITE; PS00034; ER_TARGET; UNKNOWN_1.
 SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;

Query Match 7.9%; Score 196; DB 11; Length 481;
 Best Local Similarity 21.0%; Pred. No. 9.3e-08;
 Matches 100; Conservative 71; Mismatches 194; Indels 112; Gaps 20;

QY 11 ATLFGSTGTALPRGNHSCDVTVDGYQCFPEISHLWGOYSPFFSLAD--ESA-----I 63
 Db 21 AALLSSFARCSLPRGD-----PVASVL---SPVFGTKTRYEDANPWL 62
 QY 64 SPDVPK-----GCRVTFVQVLSRHGARYPTSSSKYSALIEAIQKNATAFKGYAF 115
 Db 63 DPVAPRDPPELLAGTCTPVOLVALIRHGTTRYPTTKQIRKLKQLOGLLQRESRDGGSQA 122
 QY 116 LKTYNTLGADDLTPFGEQOMVNSGKIFRYRYKALARKIVP--FVR-----ASGSD 164
 Db 123 AALAEMPLAVGD---WMDGOLVEKGRQDMQLALRALAALFPDLFSRENYDLRLITSSKH 179
 QY 165 RVTASAEKIEGFQSAKLADPGANPHQAS-----PVINVIPEGAGYNNITLDHGLCTAF 218
 Db 180 RCVSSAAFLQGLW---QHYHPLGPLPPDVSDMECGPPRINDKL-----MRFDDH--CEKF 229
 QY 219 ESELDGDDVEANFTAVFAPPRARLEAHLPG-----VNLTDEDVNNL- 260
 Db 230 L-----TDVERNATALY-----HYEAFYTGEMQKVLKVAATLQVPMNSLNADLIQVA 278
 QY 261 MDMCPFDVARTSDATOLSPCDLFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVFN 320
 Db 279 FTCTSFDLAIKGVH---SPWCDVFDVDDARVLEYLNDLKOYWKRSYGYTINSRSCNLF 334
 QY 321 NELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKP 380
 Db 335 QDIFLHLDKAVEQKORSQ-----PVSSPVIQFQHAETLLPLLISLMGYFKDKPEP 383

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:27 ; Search time 48.78 Seconds
(without alignments)
682.417 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDVTVDYQCFETSHLW.....DFVEGLSFARSGGNWAECPA 437

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1688	72.7	467	1 JN0889	3-phytase (EC 3.1.1.3.8)
2	1684	72.6	467	1 JN0556	3-phytase (EC 3.1.1.3.8)
3	1665	71.7	441	1 JN0482	3-phytase (EC 3.1.1.3.8)
4	356	15.3	467	2 S53476	acid phosphatase (EC 3.1.1.3.8)
5	356	15.3	467	2 S48996	acid phosphatase (EC 3.1.1.3.8)
6	353	15.2	467	2 PABYCC	acid phosphatase (EC 3.1.1.3.8)
7	347	15.0	467	1 PABYCC	acid phosphatase (EC 3.1.1.3.8)
8	337	14.5	468	2 S52495	acid phosphatase (EC 3.1.1.3.8)
9	334.5	14.4	468	2 JC4285	acid phosphatase (EC 3.1.1.3.8)
10	307.5	13.2	479	1 JN0890	acid phosphatase (EC 3.1.1.3.8)
11	303.5	13.1	479	1 JN0715	acid phosphatase (EC 3.1.1.3.8)
12	298	12.8	453	2 A25326	acid phosphatase (EC 3.1.1.3.8)
13	290	12.5	463	2 S14119	acid phosphatase (EC 3.1.1.3.8)
14	285	12.3	463	2 T39929	thiamin-repressible
15	154	6.6	468	2 A86233	hypothetical protein
16	149.5	6.4	465	2 JEO369	histidine acid phosphatase
17	125	5.4	465	2 T19118	acid phosphatase (EC 3.1.1.3.8)
18	113.5	4.9	423	2 A33395	acid phosphatase (EC 3.1.1.3.8)
19	111	4.8	452	2 T20556	hypothetical protein
20	111	4.8	693	1 S61067	homocitrate hydratase
21	108.5	4.7	421	2 S14742	acid phosphatase (EC 3.1.1.3.8)
22	107.5	4.6	381	2 JH0152	acid phosphatase (EC 3.1.1.3.8)
23	106	4.6	2205	2 T08615	aggregation factor
24	105	4.5	537	2 S47770	secreted acid phosphatase
25	105	4.5	888	2 T46726	secreted acid phosphatase
26	104.5	4.5	413	2 T18945	hypothetical protein
27	104.5	4.5	423	1 S06167	acid phosphatase (EC 3.1.1.3.8)
28	104	4.5	1025	2 G81722	polymorphic membrane protein
29	102	4.4	1081	2 T51613	DNA mismatch repair

30 101 4.4 419 2 T39920
31 100 4.3 407 2 T40582
32 100 4.3 556 2 S67097
33 98.5 4.2 583 2 S46265
34 98.5 4.2 1527 2 JEO336
35 97.5 4.2 336 2 A24430
36 97 4.2 730 2 T16455
37 97 4.2 1458 2 A45665
38 96.5 4.2 844 2 T37690
39 96.5 4.2 1350 2 G36793
40 96.5 4.2 2290 1 GNNYE
41 95.5 4.1 357 2 T07851
42 95.5 4.1 394 1 KIBSGM
43 95.5 4.1 647 2 A84265
44 95.5 4.1 986 2 T33135
45 95 4.1 455 2 B71480

ALIGNMENTS

RESULT 1

JN0889
3-phytase (EC 3.1.1.3.8) A precursor - Aspergillus awamori
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phytase
C:Species: Aspergillus awamori
C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
C:Accession: JN0889
R:Piddington, C.S.; Houston, C.S.; Palohelimo, M.; Cantrell, M.; Miettinen-Oinonen, A.
Gene 133, 55-62, 1993
A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op
A:Reference number: JN0889; MUID:94040796
A:Accession: JN0889
A:Molecule type: DNA
A:Residues: 1-467 <PID>
A:Cross-references: GB:L02421; NID:q166518; PIDN:AAAL6898.1; PID:q166519
A:Experimental source: strain ALK0243
A:Note: part of the sequence, including the amino end of the mature protein, was conf
C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic
C:Genetics:
A:Gene: phytA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-467/Product: 3-phytase A #status experimental <MAT>
F:27,59,105,120,207,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 72.7%; Score 1688; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 5.2e-132;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYOC-PEISHLWGOYSPFFSLADESAISDPVPGKRVTFVQVLSRGARY 58
DB 27 NQSTCTDVGQYQCFSETSHLWGOYAPFFSLANESAISDPVPGKRVTFVQVLSRGARY 86
QY 59 PTSSSKKYSALIERIQKN-ATFKGKYAFKTYNTVLGADLTFFGENOMVNSGKIFYR 117
DB 87 PTESKGYKYSALIEETQQNVTFDQKYAFKTYNTVLGADLTFFGENOMVNSGKIFYR 146
QY 118 YKALNVPVFRASGSDRVIASAKFIEGFSQAKLADP---AQASPVINVLIPGSGY 174
DB 147 YESLRNIIPIFRSSGSRVIAEGKEIEGFSQAKLADPRAQPGSSPKIDVIVSEASS 206
QY 175 NNTLDHGLCTAFEDSTLGDDAEANFTAFAPPIARLE-ALPGVNLTDVNLMDMCPFF 233
DB 207 NNTLDPGICTVEDELADTVANFTAFAPPIARLENDLSGVLTIDTEVLYLMDMCSF 266
QY 234 DTVAITSQATQSPCDLFTADEW-QYDYLQSL-KYYGYGAGNPLGPGAGVGF-NELIAR 290
||||| : : : ||||||||| ||| ||||||| ||||||| ||||||| |||||||

probable glucanase
hypothetical prote
probable membrane
cysteine proteinas
canaliculus multia
glyceraldehyde-3-p
hypothetical prote
adult-specific bru
hypothetical prote
genome polypeptide
anakin (EC 3.4.22
phosphoglycerate k
hypothetical prote
hypothetical prote
probable agx-1 hom

Db 267 DTISTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 326

Qy 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 350

Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 386

Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEKEPLVRLVNDVRVPLHGCVD 409

Db 387 ENITQTDGFSAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDVRVPLHGCVD 439

Qy 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437

Db 440 ALGRCTRDSFVRGLSFARSGGDWAECA 467

RESULT 2

JN0656

3-Phytase (EC 3.1.3.8) A precursor - Aspergillus niger

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein

C:Species: Aspergillus niger

C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0656; S28456

R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.A.M.J.

Gene 127, 87-94, 1993

A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)

A:Reference number: JN0656; MUID:93252284

A:Accession: JN0656

A:Molecule type: DNA

A:Residues: 1-467 <N>

A:Cross-references: GB:Z16414; MID:g2392; PIDN:CAA78904.1; PID:g2393

A:Experimental source: strain NRRL3135

A:Note: parts of the sequence, including the amino end of the mature protein, were confirmed by sequencing.

C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and inorganic orthophosphate from phytate

C:Genetics:

A:Gene: phyA

A:Introns: 15/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosphatase

F:1-23/Domain: signal sequence #status predicted <Sig>

F:24-467/Product: 3-phytase A #status experimental <MAT>

F:27, 59, 105, 120, 207, 230, 339, 352, 376, 388/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:81, 361/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 72.6%; Score 1684; DB 1; Length 467;

Best Local Similarity 73.9%; Pred. No. 1.1e-131;

Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESATSIPDPKGCRTVFQVLSRHGARY 58

Db 27 NQSSCDTVDGQYQCFSETSHLWGOYAPFFSLANESVISPVPAGCRVTFPAQVLSRHGARY 86

Qy 59 PTSSKSKYSALIERIQKNA-TFKGYAFILKTYNTLIGADDLTPFGENQMVNSGIKFYR 117

Db 87 PTDSKGGKYSALIEIQONATTPDGKYAFILKTYNTSLGADDLTPFGELVNSGIKFYR 146

Qy 118 YKALARNIVPVRASGSDRVTSASEKIEGFSQSAKLADP---AHOASPVINVIPEGSY 174

Db 147 YESLIRNIVPPIRSQSGSRVTSAGKKEIEGFSQSAKLADP---AHOASPVINVIPEGSY 206

Qy 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233

Db 207 NNTLDHGLCTAFEDSELADTVEANFTATFVPSIRQLENDLSGVTLTDEVTYLMDCSF 266

Qy 234 DTIVARTSDATQLSPCDLFTADEW-OYDYLQSL-KYYGYGAGNPLGPAQGVG-NELIAR 290

Db 267 DTISTSTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 326

Qy 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 350

Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 386

Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEKEPLVRLVNDVRVPLHGCVD 409

Db 387 ENITQTDGFSAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDVRVPLHGCVD 439

Qy 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437

Db 440 ALGRCTRDSFVRGLSFARSGGDWAECA 467

RESULT 3

JN0482

3-Phytase (EC 3.1.3.8) A - Aspergillus ficuum

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein

C:Species: Aspergillus ficuum

C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

C:Accession: JN0482; PN0023

R:Ullah, A.H.J.; Dischinger Jr., H.C.

Biochem. Biophys. Res. Commun. 192, 747-753, 1993

A:Title: Aspergillus ficuum phytase: Complete primary structure elucidation by chemically modified peptide sequencing

A:Reference number: JN0482; MUID:93249451

A:Accession: JN0482

A:Molecule type: protein

A:Residues: 1-441 <ULL>

A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening residue are not cleaved by the enzyme

R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.

Biochem. Biophys. Res. Commun. 178, 45-53, 1991

A:Title: Cyclohexanedione modification of arginine at the active site of Aspergillus phytase

A:Reference number: PN0023; MUID:91298982

A:Accession: PN0023

A:Molecule type: protein

A:Residues: 48-70 <UL2>

C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phytate

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosphatase

F:4, 36, 82, 97, 184, 207, 316, 329, 353, 365/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:58, 338/Active site: Arg, His #status predicted

F:59/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 71.7%; Score 1665; DB 1; Length 441;

Best Local Similarity 73.7%; Pred. No. 3.8e-130;

Matches 328; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

Qy 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESATSIPDPKGCRTVFQVLSRHGARY 58

Db 4 NQSSCDTVDGQYQCFSETSHLWGOYAPFFSLANESVISPVPAGCRVTFPAQVLSRHGARY 63

Qy 59 PTSSKSKYSALIERIQKNA-TFKGYAFILKTYNTLIGADDLTPFGENQMVNSGIKFYR 117

Db 64 PTDSKGGKYSALIEIQONATTPDGKYAFILKTYNTSLGADDLTPFGELVNSGIKFYR 123

Qy 118 YKALARNIVPVRASGSDRVTSASEKIEGFSQSAKLADP---AHOASPVINVIPEGSY 174

Db 124 YESLIRNIVPPIRSQSGSRVTSAGKKEIEGFSQSAKLADP---AHOASPVINVIPEGSY 183

Qy 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233

Db 184 NNTLDHGLCTAFEDSELADTVEANFTATFVPSIRQLENDLSGVTLTDEVTYLMDCSF 243

Qy 234 DTIVARTSDATQLSPCDLFTADEW-OYDYLQSL-KYYGYGAGNPLGPAQGVG-NELIAR 290

Db 244 DTISTSTVDTKLSFCDLFTHDEWIHYDYLQSLKYYHGAGNPLGPTQGVGYANELIAR 303

Qy 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 350

Db 304 LTHSPVHDDTSSNHTLDSNPATFPLNATLADSHDNTWYSIFALGLYNGTKPLSTTSV 363

Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEKEPLVRLVNDVRVPLHGCVD 409

Db 364 ENITQTDGFSAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDVRVPLHGCVD 416

Qy 410 KLGRCKLDDFVEGLSFARSGGNWAE 434

acid phosphatase (EC 3.1.3.2) precursor, constitutive - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c
C:Species: *Saccharomyces cerevisiae*
C:Date: 30-Sep-1991 #sequence revision 09-Sep-1994 #text_change 05-Nov-1999
C:Accession: S48259; S45960; S05794; A25241; S44674
R:Mannhaupt, G.; Stucka, R.; Ehmele, S.; Vetter, I.; Feldmann, H..
Yeast 10, 1363-1381, 1994
A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
A:Reference number: S48255; MUID:95208357
A:Accession: S48259
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-467 <MAN>
A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55597.1; PID:g476050
A>Note: Note the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R:Eldmann, H.; Mannhaupt, G.; Schwarze, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45927
A:Accession: S45960
A:Molecule type: DNA
A:Residues: 1-467 <PE2>
A:Cross-references: EMBL:X78993; NID:g536362; PIDN:CAA85045.1; PID:g536363; GSPDB:GN0000
R:Baiba, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
Nucleic Acids Res. 12, 7721-7739, 1984
A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeast
A:Reference number: S05794; MUID:85037940
A:Accession: S05794
A:Molecule type: DNA
A:Residues: 1-218, 'MKT', 222-467 <BAJ1>
A:Cross-references: EMBL:X01080; NID:g4148; PIDN:CAA25557.1; PID:g758281
A>Note: The authors translated the codon AAT for residue 134 as ASP and TAC for residue
R:Tait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G.
Mol. Cell Biol. 6, 1855-1865, 1986
A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
A:Reference number: A93074; MUID:87064474
A:Accession: A25241
A:Molecule type: DNA
A:Residues: 1-44 <TAI>
C:Genetics:
A:Gene: SGD:PHO3; MIPS:YBR092c
A:Cross-references: SGD:S0000296; MIPS:YBR092c
A:Map position: 2R
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-467/Product: acid phosphatase, constitutive #status predicted <MAT>
F:75/Active site: His (phosphohistidine intermediate) #status predicted
F:97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (covar)
F:337/Active site: His #status predicted

Query Match 15.2%; Score 353; DB 1; Length 467;
Best Local Similarity 26.8%; Pred.No. 2.4e-21;
Matches 117; Conservative 62; Mismatches 187; Indels 70; Gaps 16;

Qy 21 GOVSPFSLADESAISDPVPKGRVTFOVLRSRHGARYPTSSKSKYSALIERIQK-NAT 79
Db 42 GGAGPVSFPDGYGISNDLPGECEMKQLMARHGERYPYTSKGATIMKTWTKLSNYTRQ 101
Qy 80 FKGYAFLEK-TYNITLGADD-----LTPF-GENOMVNNGIKFYERYKALAR 123
Db 102 FNGSLSFLNDYYEFFIRDDDDDEMTTFANSNVNLNPYTGEDMAKRHRARFFLAQGYMF 161
Qy 124 NVVVF-VRASGSDRVISAEEKFTGFOSAKIADPAHOASPVINVIPEGSGYNNITLDHGL 182
Db 162 NOTSFPIFAASSERVHPAQYFDG-----IGDOFNLSLTQTVSEAMSAGA---NTLSAGN 213
Qy 183 CTAFEDSTLGDAEANFTAVFAPPPIRARL-EALPGVNLTDVEDVVNLMDCPFDTVARTSD 241
Db 214 ACPGWDEDANDDLDKDYTTLLDDIAKKLNKENGLNLSKDNLTFLAWCAVELNAR--- 270
Qy 242 ATQLSPFCDLTAEW-OYDYLQSL-KYGGYGANPLGPAGVG-FNELIARLTHSPVQD 298

Db 271 --GYSVDICFIETDELVRYSYGQDLVSFYQDGPYQDMIRSVGANLFNATLKLLKQSETQD 328

QY 299 HTSTNHTIDSNPATPLNATLYADGSHDNTWYSIFPALGLYNGTKPLSTTSVESIETDGY 358

Db 329 -----LKWLVSFTHDITDILNLYLTAGTIDDKNNLTAEYVDFM-GNTE 369

QY 359 AASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVRLVNDVRVPLHGGCVGDKLGRCKLDD 418

Db 370 HKSWYVPQCARVYTERFC-----SNDTVRYVINADVAVPIETCTSGPGFSCIND 420

QY 419 F-----VEGLSFAR 427

Db 421 FYDYAEKRVAGTIDFLK 436

RESULT 7

PABYC

acid phosphatase (EC 3.1.3.2) repressible, precursor - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: acid phosphatase PHO5; protein YBR0814; protein YBR093c

C:Species: *Saccharomyces cerevisiae*

C:Date: 19-Feb-1984 #sequence-revision 30-Sep-1991 #text-change 12-Nov-1999

C:Accession: S05795; A38792; S48260; S45961; A00777; A38793; S41855; B25241; A25367;

R:Baizwa, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.

Nucleic Acids Res. 12, 7721-7739, 1984

A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in *y*

A:Reference number: S05794; MUID:85037940

A:Accession: S05795

A:Molecule type: DNA

A:Residues: 1-467 <BAJ>

A:Cross-references: EMBL:X01079; NID:q4162; PIDN:CAA25555.1; PID:q758282

A:Note: the authors translated the codon TAC for residue 272 as Thr

A:Accession: A38792

A:Molecule type: protein

A:Residues: 18-45 <BAJ2>

R:Manhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann, H.

Yeast 10, 1363-1381, 1994

A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.

A:Reference number: S48255; MUID:95208357

A:Accession: S48260

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-467 <MAN>

A:Cross-references: EMBL:X78993; NID:q476045; PIDN:CAA5598.1; PID:q476051

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994

R:Feldmann, H.; Manhaupt, G.; Schwarze, C.; Vetter, I.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45927

A:Accession: S45961

A:Molecule type: DNA

A:Residues: 1-467 <E2>

A:Cross-references: EMBL:X35962; NID:g536364; PIDN:CAA85046.1; PID:g536365; GSPDB:GNO

R:Arino, K.; Oshima, T.; Kubota, I.; Nakamura, N.; Mizunaga, T.; Toh-e, A.

Nucleic Acids Res. 11, 1657-1672, 1983

A:Title: The nucleotide sequence of the yeast PHO5 gene: a putative precursor of reph

A:Reference number: A00777; MUID:83168913

A:Accession: A00777

A:Molecule type: DNA

A:Residues: 1-35, 'Y', 37-129, 'G', 131-293, 'Q', 295-445, 'V', 447-461, 'DT', 464-465, 'K', 467

A:Cross-references: EMBL:V01320; NID:q4158; PIDN:CAA24630.1; PID:g4159

A:Accession: A38793

A:Molecule type: protein

A:Residues: 18-26, 'X', 28 <ARI2>

R:Meyhack, B.; Baizwa, W.; Rudolph, H.; Hinnen, A.

EMBO J. 1, 675-680, 1982

A:Title: Two yeast acid phosphatase structural genes are the result of a tandem dupli

A:Reference number: S41855; MUID:84236032

A:Accession: S41855

A:Molecule type: DNA

A:Residues: 1-30, 'T', 32-51, 'S', 53-75 <MEY>

A:Cross-references: EMBL:M24178; NID:g172156; PIDN:AAA34868.1; PID:g172157

R:Tait-Ramardt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, O.D.; Bostian, S.J.; Thill

Mol. Cell. Biol. 6, 1955-1965, 1986

A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster with

A:Reference number: A93074; MUID:87064474

A:Accession: B25241

A:Molecule type: DNA

A:Residues: 1-44 <TAI>

R: Bergman, L.W.

Mol. Cell. Biol. 6, 2298-2304, 1986

A:Title: A DNA fragment containing the upstream activator sequence determines nucleosome

A:Reference number: A25367; MUID:87064526

A:Accession: A25367

A:Molecule type: DNA

A:Residues: 1-2, 'Y', '4-43, 'T', '45-51 <BER>

R: Silve, S.; Monod, M.; Hinnen, A.; Haguenauer-Tsapis, R.

Mol. Cell. Biol. 7, 3306-3314, 1987

A:Title: The yeast acid phosphatase can enter the secretory pathway without its N-termin

A:Reference number: A27774; MUID:88038886

A:Accession: A27774

A:Molecule type: DNA

A:Residues: 1-51, 'S', '53-60 <SIL>

A:Cross-references: GB:M17306

C: Genetics:

A:Gene: SGD:PHO5; MIPS:YBR093c

A:Cross-references: SGD:S0000297; MIPS:YBR093c

A:Map position: 2R

A:Note: YBR093c

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphatidine; phosphoprotein; phosphoric monoester hydrola

F.1-17/Domain: signal sequence status predicted <SIG>

F.18-467/Product: acid phosphatase, repressible #status experimental <MAT>

F.75/Active site: His (phosphatidine intermediate) #status predicted

F.97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (coval

F.333/Active site: His #status predicted

Query Match 15.0%; Score 347; DB 1; Length 467;
Best Local Similarity 26.5%; Pred. No. 7.5e-21;
Matches 117; Conservative 57; Mismatches 187; Indels 80; Gaps 18;

QY 21 GOYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTSSKSKYSALIERIOK-NAT 79

Db 42 GGAGPYSPGDDYIGSRDLPECEMKQLQWGRGERIPTVSLARTIKSTWYKLSNYTRQ 101

QY 80 FKGVAFLEK-TYNYTLGND-----LTPF-GENQMVNSGKIFRYRYKALAR 123

Db 102 FNGSLFNDYEFFIRDDDDLEMETTFANSDDVNLNPTGEMNAKRHARDPLAQGYAVE 161

QY 124 NIVPF-VRASGDRVIAAEKIEGFSQAKLADPAHQASPVINII-----PEGSGYNT 177

Db 162 NQTSFAVETSNKRCHDTAQYFIDG-----LGDO-----FNITQTVSEAESAGANTL 209

QY 178 LDHGLCTAFEDSTLGGDAEANTAVFAPPRIARL-EALPGVNLTDDEVNLMDCMCPDVT 236

Db 210 SACNSCPAW-DYDANDDIVNEYDTTLDIAKRLNKENKGLNTSTDASTLFSWCAFEV- 267

QY 237 ARTSDATQLSPFCDLETADW-QYDYLQSL-KYGYGAGNPLGPAQGVG-FNELIARLTH 293

Db 268 -----NAKGYSDVCDIFTKDELHYHYQQDLHTYHCGPGYDIITKSGNSLNFASVLLKQ 323

QY 294 SPVQDHTSNHTLDSNPATFPLNATLYADFSDNTMVSIFFAFGKPLSTTSVESTI 353

Db 324 SETQDQ-----KWLVSFTHTDILNLTITAGLIDDKNNLTAEVVPFM 365

QY 354 ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGGGVKGLGR 413

Db 366 -GNTFHRSWVPGGARVYTEKFC-----SNDTYVRYVINDAVVPIETCSTPGGFS 415

QY 414 CKLDDE-----VGLSFPAR 427

Db 416 CEINDFYAEKRVAGTDFLK 436

RESULT 8

S52495

acid phosphatase homolog YDL024c - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein D2815
C:Species: Saccharomyces cerevisiae
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C:Accession: S52495; S67556
R:Andre, B.; Visser, S.; Urrestazu, L.
submitted to the EMBL Data Library, February 1995
A:Description: The sequence of a 42 kb segment located on the left arm of chromosome
A:Reference number: S52492
A:Accession: S52495
A:Molecule type: DNA
A:Residues: 1-468 <AND>
A:Cross-references: EMBL:Z48432; NID:9683669; PIDN:CAA88335.1; PID:9683673
R:Urrestazu, L.A.; Andre, B.; Visser, S.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67533
A:Accession: S67556
A:Molecule type: DNA
A:Residues: 1-468 <URR>
A:Cross-references: EMBL:Z74072; NID:91430996; PIDN:CAA98583.1; PID:91430997; MIPS:YD
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 4L
C:Superfamily: yeast acid phosphatase

Query Match 14.5%; Score 337; DB 2; Length 468;
Best Local Similarity 26.0%; Pred. No. 5e-20;
Matches 119; Conservative 60; Mismatches 161; Indels 118; Gaps 19;

QY 21 GOYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYPTSSKSKYSALIERIOK-NAT 80

Db 43 GGSADYFSPFANYGPTDIPGCRLLTQVMIGRHGRIYPTSEAKD---IFEVWKISNY 99

QY 81 KGKY----AFKLT-YNYTL-----GADLLTPF-GENQMVNSGKIFRYRYKA 120

Db 100 TKYBGSLSFLNNGYEFPIDESLLEMETTIONSIDVNLNPTGEMNAKRHAREFLAKYKG 159

QY 121 LARNIVPF-VRASGDRVIAAEKIEGFSQAKLADPAHQASPVINVIPEGSGYNTLTD 179

Db 160 LMENCTNPFITTSKRIYDTAQYFAEL-----GDGFNLSLQ 197

QY 180 HGLCTAFEDSTLGG-----DDAEANTAVFAPPRIARL-EALPGVNL 220

Db 198 ----TLSENSSGANTLAAKSSCPNWSNANNILMSYSDYLENISDRNDENKGLNLS 253

QY 221 DEVVNLMDCMCPFTVARTSDATQLSPFCDLETADW-QYDYLQSL-KYGYGAGNPLGP 278

Db 254 RKDAALFSCAFEL-----NAKGSNIDIFSAELIHYSETDLTSFYQNGPGYKLIK 308

QY 279 AQGVG-FN---ELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFSDNTMVSIF 334

Db 309 SIGANLENATVKLIQSAH-----LDQKWLVSFTHTDILNLT 347

QY 335 ALGYNGTKPLSTTSVESTIDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRV 394

Db 348 TAGLIDDTLRLTNHV-PFRDHSYHRSWYIPOGARVYTEKFC-----SNDSYVRY 397

QY 395 LVNDRVPLHGGGVKGLGRCKLDDFVE-----GLSF 425

Db 398 VNDVAVPIESCSSGPGFCEGTFEYAKDLRLGVSF 435

RESULT 9

JC4285

acid phosphatase (EC 3.1.3.2) precursor - yeast (Pichia pastoris)

N:Alternate names: orthophosphoric monoester phosphohydrolase (acid optimum); Pho

C:Species: Pichia pastoris

C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999

C:Accession: JC4285

R:Payne, W.B.; Gannon, P.M.; Kaiser, C.A.

Gene 163, 19-26, 1995

A:Title: An inducible acid phosphatase from the yeast Pichia pastoris: Characterizati

Sat Oct 27 15:25:43 2001

C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-479/Product: 3-phytase #status predicted <MAT>
F:81.337/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted
F:106.191,227,250,315,340,425,442,458/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 13.2%; Score 307.5; DB 1; Length 479;

Best Local Similarity 27.4%; Pred. No. 1.5e-17;
Matches 124; Conservative 55; Mismatches 175; Indels 99; Gaps 23;

QY 9 DGYQCPETISHLWGQ--YSPFFSLADESAISDPVPGCRVTFVQVLSRHGARYPTSSSKK 66

DB 41 DGYSI--LKHGNGPYSERVSY---GIARDPPTSCEVDQVIMVKRHGERYPSPSAGKD 94

QY 67 YSALIERIOK-NAT-FKGYAFKTYNYTL-----GADDLT-PF-GENQMVNSGIKEYR 116

DB 95 IEALAKVYSINTTEYKGLAFNLDMWTYVPNECYNAETTSQPYAGLLDAYNHGNDYKA 154

QY 117 RYKAL--ARNIVPFVRASGSDRVASAEKFIQFQSAKLADPAHQASPVINVIPEGSGY 174

DB 155 RYGLMNGETVVPFF--SSGYGRVETARKFGSGF-----FGY 190

QY 175 NNTLDHGLCTAFEDSTLGDDAE-----NFTAVTAPPPIR---ARLEAL-PGV 217

DB 191 NYSNAALNIISEVMGADSLTPTCDTNDQTTCDNLTYQL-POFKVAARLNSONPOM 249

QY 218 NLTDDEVNLMDCPFDVARTSDATQLSPCDLFTADEW-QYDYLOSLEYGYGAG--- 273

DB 250 NLTAADSYNLMWASFEINAR-----PFSWNAFTQDEWVSFGYVEDLNY--YCAGPGD 303

QY 274 NPLGPAQGVGFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADESHDNTWVSIF 333

DB 304 KMAAVGANVYANASLTLLNQPKK-----AGSLFFNFHAHDTNTPIL 345

QY 334 FALGLY--NGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMQCAGGGEKEPL 391

DB 346 AALGVLPINEDLPDRVAF-----GNPYSIGNIVPMGCHLTIERLSQATALSDEG---TY 398

QY 392 VRVLVNDVRVPLHGGVDKLGRCCKLDFFVEGLS 424

DB 399 VRVLVNEAVLPFNDCTSGPGYSCPLANVTSTLN 431

RESULT 11

JN0715

3-phytase (EC 3.1.3.8) B precursor - Aspergillus ficuum

N:Alternate names: pH 2.5-Optimum acid phosphatase

C:Species: Aspergillus ficuum

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999

C:Accession: JN0715; PNO594; PNO460

R:Ullah, A.H.J.; Dischinger Jr., H.C.; Ullah, A.H.J.

Biochem. Biophys. Res. Commun. 192, 754-759, 1993

A:Title: Identification and cloning of a second phytase gene (phyB) from Aspergillus

A:Reference number: JN0715; MUID:93371452

A:Accession: JN0715

A:Molecule type: DNA

A:Residues: 1-479 <EHR>

A:Cross-references: GB:L20567

A:Accession: PNO594

A:Molecule type: protein

A:Residues: 20-101;133-146;376-399 <EH2>

R:Ullah, A.H.J.; Dischinger Jr., H.C.

Biochem. Biophys. Res. Commun. 192, 754-759, 1993

A:Title: Identification of active-site residues in Aspergillus ficuum extracellular p

A:Reference number: PNO460; MUID:93249452

A:Accession: PNO460

A:Molecule type: protein

A:Residues: 65-66,68-93 <ULL>

C:Comment: This enzyme, previously characterized as an acid phosphatase (EC 3.1.3.2),

C:Genetics:

A:Reference number: JC4285; MUID:96001238

A:Accession: JC4285

A:Molecule type: DNA

A:Residues: 1-468 <PAY>

A:Cross-references: GB:U28658; NID:9881955; PIDN:AAA85503.1; PID:9881956

A:Experimental source: GS115

C:Genetics:

A:Gene: phol

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrola

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-468/Product: acid phosphatase #status predicted <MAT>

F:84/Active site: His (phosphohistidine intermediate) #status predicted

F:163,196,256,321,360,435/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:345/Active site: His #status predicted

Query Match 14.4%; Score 334.5; DB 2; Length 468;

Best Local Similarity 27.0%; Pred. No. 8.1e-20;

Matches 124; Conservative 65; Mismatches 187; Indels 83; Gaps 22;

QY 7 TVDGYQCPETISHLWGQYSPF-----FSLADESAISDPVPGCRVTFVQVLSRHGARYPTS 61

DB 40 TDDQYNI--LRHL-GGLGPGYGYNGWGAASEI-----ESCTIDQAHLLMRHGERYPST 91

QY 62 SKSKYSALIERIOKNATFK---GRYAFKTYNYTLG-----ADLTP-----FGENQMVNS 110

DB 92 NVCKQLEALYQKL-LDADVEVPTGLSFQDYDFVSDAAWVEQETTKGFYSGLNTAFDF 150

QY 111 GIKFYRYKALARNIVP-----FVRASGSDRVTSAEKFIQFQSAKLADPAHQASPVINV 166

DB 151 GTTLREYDHLINTSEEGKLSWAGSQERVVDYTKYAQGFKNKYTD-----WVEV 203

QY 167 IIPF--GSGYNTLDHGLCTAFEDSTLGDDAEANFTAVFAPPPIRARLEAL-PGVNLDE 222

DB 204 VALEEKSGSLNSTARISCPNVNSHIYD---GDFPNDAIERADRLNTLSPGFNITAD 260

QY 223 DVYNLMDMCPFDVARTSDATQLSPCDLFTADEWQYD-YLQSLK-YGYGAGNPLGPAQ 280

DB 261 DIFTIALYCGFELNVRGE-----SSPCDLSREALLYTAYLRDLGWYNNVNGNPLGKTI 315

QY 281 GVGFNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADESHDNTWVSIFFAFGLYN 340

DB 316 GYVYANATROLLEN-----TEADPRDP-----LYFSFSDTDLLOVFTSLGLFN 360

QY 341 GTK-PLS---TTSVESIETDGYAASWTVPFAARAYVEMMQCAGGGEKEPLRVVL 395

DB 361 VTDLPDQIQFQTSFKSTE-----IVPMGARLLTERLLCTV-----EGEEKYVVRTI 407

QY 396 VNRVVPVPHGGVDKLGRCCKLDFFVEGLSFARSGNNAE 434

DB 408 LNDVAVPLSDCSGPGFSCPLNDYVSRLEALNEDSDFAE 446

RESULT 10

JN0890

acid phosphatase (EC 3.1.3.2) precursor - Aspergillus awamori

C:Species: Aspergillus awamori

C:Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0890

R:Pidington, C.S.; Houston, C.S.; Palohelmo, M.; Cantrell, M.; Miettinen-Oinonen, A.; N

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-opti

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0890

A:Molecule type: DNA

A:Residues: 1-479 <PID>

A:Cross-references: GB:L02420; NID:g166481; PIDN:AAA16897.1; PID:g166482

A:Experimental source: strain ALK0243

C:Comment: The highly similar enzyme from A. ficuum has been shown to have 3-phytase (EC

C:Genetics:

A:Gene: aph

A:Introns: 261/1; 300/2; 335/2

Sat Oct 27 15:25:43 2001

Db 212 VFKNDFHKNATDAAHAYVRNIFIEPIVNRLLAKYFDSSYKLTINDVRSIFYCEYEIAK 271
 QY 239 TSDATOLSPFCDLFTADAW-QYDYLOSLKYYGYGAGNPLGPAQGVGFENELIARLTHSPVQ 297
 Db 272 DH-----SDFCSIFTPSEFLNFEYDSDL-QAYGG-----PVS 304
 QY 298 DHTST-----NHTLDS--NPATFPLNATLYADFSDHNTMWSIFFALGLY---NKGKPLS 346
 Db 305 EWASTLGGAYINLADSLRNVTNPDFDRKVFLLAFTHDSNLIIPVEAALGFDPDITPONPLP 364
 QY 347 TTSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGGEKEPEPLRVLVNDRVPLHCC 406
 Db 365 TD--KNLYTSQKTSFVFPAGNLITELFEC-----SDSKYYVRHLVNOOVYPLIDC 414
 QY 407 GVDKLR-----CKLDDFVGLSFGSGGNWAEFC 436
 Db 415 GYGPSGTSDDLGLCELQALYNSPIRANSTNSGISTF 448

RESULT 15

A86233

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A86233

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzia,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A86233

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 <S>T0>

A:Cross-references: GB:AF005172; NID:g2160177; PIDN:AAB60740.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 6.6%; Score 154; DB 2; Length 468;

Best Local Similarity 22.9%; Pred. No. 7.4e-05;

Matches 96; Conservative 52; Mismatches 184; Indels 88; Gaps 18;

QY 38 DVPKGRVTFVQVLSRHGARYPTSSKKKYSALIERI-----Q 75

Db 50 NVPSECTPIHLNLVARHGTSPTRKRRLRELESAGRFKELVRDAEARKLPDCKIPGWLQ 109

QY 76 KNATFGKYAFKTYNTYLCADDLTPFGENOMVNSGKIFYRYKALARN-----IVPFVRA 131

Db 110 WSPWEGK-----VKGGLIRIGEDELQYLGIRVRFRFSLFEEDYHPDYTTIRA 159

QY 132 SGSDRVIAAEKIEGFSQAKLADPAHQASPVINVIIEGSGYNNLTHGLCTAPEDSTL 191

Db 160 TQIPRASASAVAFGMGLFSEK-----GNL-----GPGNRFAF-----AVTSEN 198

QY 192 GDAAENAFVAPPFIRARLEALPGVNLTDVNLMDMC-----PFTVARTSDATQLSPF 248

Db 199 ASDTKLRFEECCONKYSYKAKEPAVDKLEPVNLKITASVAKRYDLKFTKODISLWFL 258

QY 249 CDLFTADEWQYDYLOSLKYYGYGAGNPLGPAQGVGFENELIARLTHSPVODHTSTNHTLDS 308

Db 259 CKQVALLWT-DDLEVLFLKGY--GNSLNKMGV---PLEDVLHSMEEAIAKREKLP- 311

QY 309 NPATFPLNATLYADFSDHNTMWSIFFALGLY-NGT-----KPLSTTSVESIETDGYA 359

Db 312 -PGSYE-KARL--RFAHAETIVFSCLLGLFDGSEFEKIOREKPLELPP-OPPKTRDR 366

QY 23 YSPFSLADESAISPDPKGRVTFVQVLSRHGARYPTSS-----KSKKYSALIERIQK 76
 Db 43 HEPYF-----NGPTTSPFESCAIKQVHLLQHGSRNPTGDDTATDVSAQYIDIFONKLL 97
 QY 77 NATFGKYA-----FLKTYNTYL--GADDLTPFGENOMVNSGKIFYRYKALARNIV 126
 Db 98 NGSIPIVNSYPENPLFYFKHHTVPVKAENADOLSSGRIELFDLGRQVFEERYELFDTDV 157
 QY 127 PVRASGSDRVIAAEKIEGFSQAKLADPAHQASPVINVIIE--SGYNNLTHGLCT 184
 Db 158 YDINTAAQERVVDSAEWESYGMFGDMQKNTN-----FIVLPEDDSAGANSLAMTSC 211
 QY 185 AFEDSTLGDG-----AEANFTAVFAPPFIRARLEAL--PGVNLTDVNLMDMCPTDVAR 238
 Db 212 VYEDNNIDENTTAAHTSWRNVLKPIANRLNKYFDSYGLNTVSDVRSLLYICVVEITLR 271
 QY 239 TSDATOLSPFCDLFTADAW-QYDYLOSLKYYGYGAGNPLGPAQ-----GVGFNELIAR 290
 Db 272 DN-----SDFCSIFTPSEFLNFEYDSDLQAYWG-----GPASEWASTLGGAYVNNLANN 321
 QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSDHNTMWSIFFALGLY---NKGKPLST 347
 Db 322 L-----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFDPDITPEHPLPT 365
 QY 348 TSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGGEKEPEPLRVLVNDRVPLHCC 407
 Db 366 D--KNLYTSQKTSFVFPAGNLITELFEC-----SDKNYYVRHLVNOOVYPLIDCG 415
 QY 408 VDKLR-----CKLDDFVGLSFGSGGN 431
 Db 416 YGPGSAGDGLCELSAYLNSVRVNSTN 443

RESULT 14

T39929

thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T39929

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.

submitted to the EMBL data Library, May 1998

A:Reference number: Z21857

A:Accession: T39929

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-463 <LYN>

A:Cross-references: EMBL:AL023286; PIDN:CAAL18863.1; GSPDB:GN00067; SPDB:SPBC21H7.03c

A:Experimental source: strain 972h; cosmid c21H7

C:Genetics:

A:Gene: SPDB:SPBC21H7.03c

A:Map position: 2

C:Superfamily: yeast acid phosphatase

Query Match 12.3%; Score 285; DB 2; Length 463;

Best Local Similarity 24.9%; Pred. No. 1e-15;

Matches 113; Conservative 69; Mismatches 184; Indels 88; Gaps 19;

QY 23 YSPFSLADESAISPDPKGRVTFVQVLSRHGARYPTSS--SKSKKYS-----L 70

Db 43 HEPYFDGLDSA-----FPETCEIQVHLLQHGSRNPTGDTATDVYSSQYLNNFQEKLL 97

QY 71 IERIQKATF-KGKYAFKTYNTYL--GADDLTPFGENOMVNSGKIFYRYKALARNIV 126

Db 98 NGSIPIVNSYPENPLCYFKHHTVPVDAENADOLSSRGLFLDGRQYKRYKLFDSYV 157

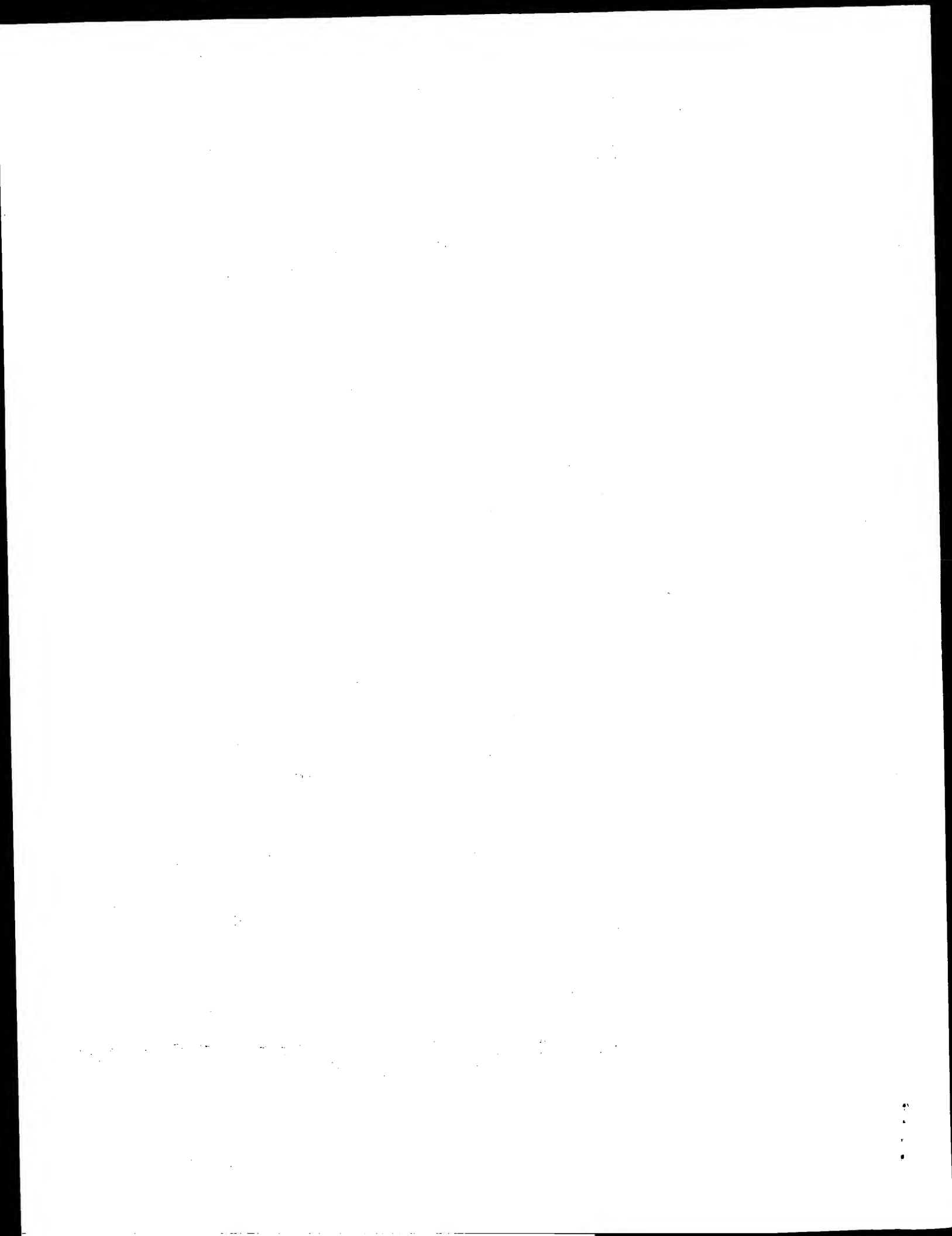
QY 127 PVRASGSDRVIAAEKIEGFSQAKLADPAHQASPVINVIIEP--SGYNNLTHGLCT 184

Db 158 YDINTAEQERVVESAKWTFYGLFGDKMEKTN-----FIIISEGKAAGANSLMYNACP 211

QY 185 AFEDSTL-----GDDAEANFTAVFAPPFIRARLEAL--PGVNLTDVNLMDMCPTDVAR 238

Oy 360 ASWTVPFAARAYVENMOCEAGGGGEGEKEPLVRVLVNDRVVPLHGCGVDKLGRCCKLDDF 419
Db 367 GSTMAPFGGNNILVLYSCPA-----ESSPKYFVQVLHNEHPIAVPGC--DGKDFCPLDF 419

Search completed: October 26, 2001, 15:10:28
Job time: 6334 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:18:02 ; Search time 29.88 seconds
(without alignments)
500.992 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NNSCDTVDGQCPEISHLW.....DFVEGLSFARSGNWAECFA 437

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1688	72.7	467	1	PHYA_ASPAW
2	1684	72.6	467	1	PHYA_ASPNG
3	1676.5	72.2	463	1	PHYB_EMENI
4	356	15.3	467	1	PPAB_YEAST
5	356	15.3	467	1	PPAC_YEAST
6	353	15.2	467	1	PPA3_YEAST
7	347	15.0	467	1	PPA5_YEAST
8	337	14.5	468	1	PPAD_YEAST
9	334.5	14.4	468	1	PPAI_PICPA
10	307.5	13.2	479	1	PHYB_ASPAW
11	305.5	13.2	479	1	PHYB_ASPNG
12	300.5	12.9	469	1	PPA5_KLULA
13	298	12.8	453	1	PPAL_SCHPO
14	290	12.5	463	1	PPA2_SCHPO
15	125	5.4	755	1	PPAX_CAEEL
16	113.5	4.9	423	1	PPAL_RAT
17	111	4.8	693	1	LYS4_YEAST
18	108.5	4.7	421	1	PPAL_MOUSE
19	107.5	4.6	381	1	PPAP_RAT
20	104.5	4.5	413	1	PPAW_CAEEL
21	104.5	4.5	423	1	PPAL_HUMAN
22	102	4.4	1081	1	MSH3_ARATH
23	98.5	4.2	341	1	AAPJ_RHILV
24	98.5	4.2	583	1	CYSP_PLAVI
25	98.5	4.2	1527	1	MRP3_HUMAN
26	97.5	4.2	392	1	G3PA_TOBAC
27	97	4.2	1458	1	PHLX_RABIT
28	96.5	4.2	844	1	YD25_SCHPO
29	96.5	4.2	1350	1	VG72_HSVII
30	96.5	4.2	2290	1	POLG_EMCV
31	95.5	4.1	394	1	PGK_BACME
32	94	4.0	776	1	ISOA_PSEAY
33	93.5	4.0	776	1	ISOA_PSESP

34	93	4.0	1047	1	RIR1_CHLTR
35	92.5	4.0	976	1	AMYBUTETI
36	92	4.0	551	1	AMYB_THETU
37	91.5	3.9	296	1	NPL_ECOLI
38	91.5	3.9	413	1	AGP_ECOLI
39	91	3.9	1052	1	MGPC_MYCGE
40	91	3.9	1522	1	MRP3_RAT
41	90.5	3.9	602	1	PEHX_ERWCH
42	90.5	3.9	2193	1	POLG_HE7IB
43	90	3.9	356	1	QUEA_ECOLI
44	89.5	3.9	1055	1	RPOC_PEDAC
45	89	3.8	4289	1	TENX_HUMAN

O84834 chlamydia t
P30269 butyrivibri
P19584 thermoanaer
P06995 escherichia
P19926 escherichia
P22747 mycoplasma
O88563 rattus norv
P15922 erwinia chr
Q66478 human enter
P21516 escherichia
P77917 pediococcus
P22105 homo sapien

ALIGNMENTS

RESULT 1
PHYA_ASPAW
ID PHYA_ASPAW STANDARD; PRT; 467 AA.
AC P34753;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 3-PHYTASE A (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL-HEXAPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA OR PHY.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=105351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALK0243;
RA MEDLINE=94040796; Pubmed=8224894;
RA Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
RA Miettinen-Oinonen A., Nevalainen H., Rambosek J.A.;
RT "The cloning and sequencing of the genes encoding phytase (phy) and
RT pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
RT awamori."
RL Gene 133:55-62(1993).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L02421; AAA16898.1; -
CC PIR: JN0889; JN0889.
CC HSSP: P34752; IHP.
CC InterPro: IPR000560; -
CC Pfam: PF00328; acid_phosphat; 1.
CC PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC Hydrolase; Glycoprotein; Signal.
CC SIGNAL 1 23 POTENTIAL A.
CC CHAIN 24 467 3-PHYTASE A.
CC ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
CC ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
CC DISULFID 31 40 BY SIMILARITY.
CC DISULFID 71 414 BY SIMILARITY.
CC DISULFID 215 465 BY SIMILARITY.
CC DISULFID 254 282 BY SIMILARITY.
CC DISULFID 436 444 BY SIMILARITY.

van Loon A.P.;
"Cloning of the phytases from Emericella nidulans and the thermophilic fungus Talaromyces thermophilus". Biochim. Biophys. Acta 1353:217-223(1997).
-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE.
-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; U59803; AAB96871.1; -.
InterPro; IPR000560; -.
Pfam; PF00328; acid_phosphat_1.
PROSITE; PS00616; HIS_ACID_PHOSPAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPAT_2; 1.
Hydrolase; Glycoprotein; Signal.
SIGNAL 1
CHAIN 20 463
ACT_SITE 80 80

(BY SIMILARITY).
NUCLEOPHYLIC ACCEPTOR (BY SIMILARITY).
PROTON DONOR (BY SIMILARITY).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
ECC5827DI1C82A2 CRC64;

FY Match 72.2%; Score 1676.5; DB 1; Length 463;
tLocal similarity 73.4%; Pred. No. 1.4e-131;
ches 326; Conservative 43; Mismatches 58; Indels 17; Gaps 10;

2 SHSCDTVD-GYOC-PEISHLWCOXPFFSLADESAISPDPKGCRCVTFVOVLSRHGARYP 59.
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
26 NISCNTADGGYCQFPNVSHVWGVSYPYSIFESAISDPVHGCEVTFVOVLSRHGARYP 85.
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 TSCKSKYSALIERIQKNAT-FKGKYAPFKTYNTLGADDLTPEGNQMVSNGIKFYRRY 118.
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 TESKSKASGLTEAQKNATSFWGQYAFLESYNTLGADDLTIFGENQMWDSCAKFYRY 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 KALARNIVPFVRAGSDRVIASAEKFTIEGQSACLAD-PAHQASPVINVIIPGSGYNNT 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
146 KNLRKNTPTFRAGSDRDVVASAEKFTNGRFKAQLHDHGSKRATPVNVNIPEIDGFNT 205
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 LDHGLCTAFESTLDGDAENFTAVFAPIRARLE-ALPGVNLTDEDVVNLMDKCFDTV 236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 LDHSPTCVSFENDERADELEANFTAIMGPPIRKRLNDLPKIKUTNEVILYLMDSFDTM 265
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 ARTSDATOLSPCDLFTADEW-QDYILOS-L-KYYGYAGNPLGPAOGVG-FNELIARLTQ 325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 ARTANGELSPFCALITEKEWLQDYVLQSLSKYGYGAGSLPGPAQGIGFTNELIARLTQ 293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
294 SPVDHTSTNHTLDSNPATFLNATLYADFSDHNTVMYSTFEALGLYNGTKPLSTTSVESI 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 SPVDQNSTNHTLDSNPATFLDKRLYADFSDHNSMISIFFAMGLYNGTQPLSMDSVESI 385
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 -ETDGYAASVTVPFAARAYEMMQCEAGGGGEKEPILRVLVNDRVVPVPLHGCCVDKL 412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 QEMDGYAASVTVPFGARAYELMOCE-----KKPEPLRVLVNDRVVPVPLHGAVDK 437
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

119 KALARNIVPVRASGSDRVIAAEKFIEGFSQAKLAD-PAHQASPVINVTIPEGSYNNT 177
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
146 KNLARKNTPFTRASGSDRVVSAEKFINGFRKAQLDHGSKRATPVVVNIIPIDGFNT 205
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
178 LDHGUCTAFEDSTLLGDAAENFTAVFAPPTRARLE-ALPGVNITDEVDVNLMDMCPFDTV 236
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
206 LQHSICVSFENDERADELEANFTAIMGPPIRKLENDLPQIKLTENVYLVLMDCMSFDTM 265
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
237 ARTSATOLSPFCDLFTADEW-QVDYLOSL-KYYGYAGNPLGPAGGVCF-NELLARTHT 293
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
266 ARTAHGTLESFPCAIFTEKEWLQDYVLQSLSKYYGYAGSPLGPAGQGIGFTNELLARLTQ 325
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
294 SPVODHSTNHTLDSNPATPELNATLYADFSDHNTMVSTIFFAIGLYNGTKPLSTTSVESI 353
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
326 SPVODNTSTNHTLDSNPATPELDKRILYADFSDHNSMISIFFAMGLYNGTQPLSMDSVESI 385
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
354 -ETDYGAASTVPFAARAYEMMOCEAGGGGEKEPLRVRLVNDRVVPHGCGVDKL 412
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
386 QEMDYGAASTVPFGARAYELMOCE-----RKPEPLRVRLVNDRVVPHGCAVDKFG 437
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

FT	17	17	A -> L (IN REF. 2).
CONFLICT	82	83	VS -> AR (IN REF. 2).
CONFLICT	150	150	R -> H (IN REF. 2).
CONFLICT	354	354	K -> Q (IN REF. 2).
CONFLICT	423	423	D -> G (IN REF. 2).
SEQUENCE	467 AA;	52757 MW;	AECDCLC046B36C3 CRC64;
Query Match	15.3%;	Score 356;	DB 1; Length 467;
Best Local Similarity	26.9%;	Pred. No. 5.2e-22;	
Matches 119;	Conservative	63;	Mismatches 190; Indels 70; Gaps
QY	15	ETSHLWGQVSPFFSLADESATSPDPVKGCRVTFVOVLSRRHGARYPTSSSKYSALIERI 74	
DB	15	ETSHLWGQVSPFFSLADESATSPDPVKGCRVTFVOVLSRRHGARYPTSSSKYSALIERI 74	
QY	36	EIPFLGGSGPYPCDYSRDLRPLSCCEMKQVMYGRGERYPVSVRAKSLMTTWYKL 95	
DB	36	EIPFLGGSGPYPCDYSRDLRPLSCCEMKQVMYGRGERYPVSVRAKSLMTTWYKL 95	
QY	75	QK-NATFKGKYAFLK-----TNYNLT-----GADDLTFE-GENQMVNSGIKFYR 117	
DB	75	QK-NATFKGKYAFLK-----TNYNLT-----GADDLTFE-GENQMVNSGIKFYR 117	
QY	96	SNVTGQFSALSFLLNDYEFFTRDTKMLEMTTILANSVNVLPNPTGEMNAKRHARDFLAQ 155	
DB	96	SNVTGQFSALSFLLNDYEFFTRDTKMLEMTTILANSVNVLPNPTGEMNAKRHARDFLAQ 155	
QY	118	YKALARNIVPF-VRASGSDRVITASKEKIEFGQSAKLADPAHOASPVINVIIEGSGYNN 176	
DB	118	YKALARNIVPF-VRASGSDRVITASKEKIEFGQSAKLADPAHOASPVINVIIEGSGYNN 176	
QY	156	YGVWVENQTSFAVTSNSRCHDTAQYFDG-----LCDKENISLQTISE-AESAGANT 208	
DB	156	YGVWVENQTSFAVTSNSRCHDTAQYFDG-----LCDKENISLQTISE-AESAGANT 208	
QY	177	TLDHGGLCTAFEDSTGLDGAENFTAVFAPPIRARL-EALPGVNLTDEEDVNLMDMCFDT 235	
DB	177	TLDHGGLCTAFEDSTGLDGAENFTAVFAPPIRARL-EALPGVNLTDEEDVNLMDMCFDT 235	
QY	209	LSAHGHPAWDDD-VNDDILKVDYTKYLSGIAKRLNKENKGLNLTSSDNTFFAWCAYEI 267	
DB	209	LSAHGHPAWDDD-VNDDILKVDYTKYLSGIAKRLNKENKGLNLTSSDNTFFAWCAYEI 267	
QY	236	VARTSDATQLSPFCDLFTADEW-QYDYQLSLK-YVYGAGNPLGPAOGVG-FNELIARLT 292	
DB	236	VARTSDATQLSPFCDLFTADEW-QYDYQLSLK-YVYGAGNPLGPAOGVG-FNELIARLT 292	
QY	268	NAR-----GYSDICNFTKDELVRFSYGODLFTYQYTGPGYDVVRSGCANLFNASVKLLK 322	
DB	268	NAR-----GYSDICNFTKDELVRFSYGODLFTYQYTGPGYDVVRSGCANLFNASVKLLK 322	
QY	293	HSPVQDHTSNHTLDSNPATFPLNATLYADFSDHNTMVSIEFALGLYNGTKPLSTTSVES 352	
DB	293	HSPVQDHTSNHTLDSNPATFPLNATLYADFSDHNTMVSIEFALGLYNGTKPLSTTSVES 352	
QY	323	ESEVQDQ-----KWLFSFTHDITLNLVTLTGIIIDKKNLTAEHVPF 364	
DB	323	ESEVQDQ-----KWLFSFTHDITLNLVTLTGIIIDKKNLTAEHVPF 364	
QY	353	IETDGYASWTVPFAARAYVEMMQCBAGGGGEGEKEPLRVLVNDVRVPLHGGGVDKLG 412	
DB	353	IETDGYASWTVPFAARAYVEMMQCBAGGGGEGEKEPLRVLVNDVRVPLHGGGVDKLG 412	
QY	365	ME-NTFHRWVVPQGARVYTEKQC-----SNDTVRVVINDAVVPIETCTGTGPGF 414	
DB	365	ME-NTFHRWVVPQGARVYTEKQC-----SNDTVRVVINDAVVPIETCTGTGPGF 414	
QY	413	RCKLDDF-----VEGLSFAR 427	
DB	413	RCKLDDF-----VEGLSFAR 427	
QY	415	SCEINDFYDYAEKRVAGTDFLIK 436	
DB	415	SCEINDFYDYAEKRVAGTDFLIK 436	
RESULT 5			
PPAC_YEAST	STANDARD;	PRT;	467 AA.
ID	PPAC_YEAST		
AC	P38693;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	01-FEB-1995 (Rel. 31, Last annotation update)		
DE	ACID PHOSPHATASE PHO12 PRECURSOR (EC 3.1.3.2).		
GN	PHO12 OR YHR215W.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRALN-S288C / AB972;		
FX	MEDLINE=94378003; PubMed=8091229;		
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,		
RA	Du Z., Favellio A., Fulton L., Gattung S., Geisel C., Kirsten J.,		
RA	Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,		
RA	Latreille P., Louis E.J., Macri C., Mardis E., Manes S., Mouser L.,		
RA	Nhan M., Rifkin L., Riles L., St Peter H., Trevasik E., Vaughan K.,		
RA	Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,		
RA	vaudin M.;		
RA	"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome		
RT	VIII.";		
RT	Science 265:2077-2082(1994).		
CC	-!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN		

QY 413 RCKLDDFVEGLSFARSGGNWAECF 436
 || ||||:||||:||||||| ||
 DB 438 RCTDDWVEGLNFARSGGNWKTCP 461

RESULT 4

PPAB_YEAST	STANDARD;	PRT;	467 AA.
ID P35842;			
AC 01-JUN-1994 (Rel. 29, Created)			
DT 01-JUN-1994 (Rel. 29, last sequence update)			
DT 01-NOV-1997 (Rel. 35, last annotation update)			
DE ACID PHOSPHATASE PHO11 PRECURSOR (EC 3.1.3.2) (P56).			
DT PHO11 OR YAR071W.			
GN Saccharomyces cerevisiae (Baker's yeast).			
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX NCBI_TaxID=4932;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=S288C / AB972;			
RA MEDLINE=95249563; PubMed=7731988;			
RX Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,			
RT Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,			
RA Storms R.K.;			
RA "The nucleotide sequence of chromosome I from Saccharomyces			
RT cerevisiae.";			
RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).			
[2]			
RN SEQUENCE FROM N.A.			
RP Chen J.Y., Gong Y.I., Ao S.Z.;			
RA "The primary structure of acid phosphatase gene PHO11 in S.			
RT cerevisiae and comparison with other gene families.";			
RL Acta Biochim. Biophys. Sin. 21:437-444(1989).			
CC -! CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN			
CC ALCOHOL + ORTHOPHOSPHATE.			
CC -! INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS			
CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.			
CC -! PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.			
CC -! SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.			
CC			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC			
DR EMBL: L28920; AAC09508.1; -			
DR EMBL: U19789; AAA73479.1; -			
DR HSSP: P34752; 1IHP.			
DR SGD: S0000094; PHO11.			
DR InterPro: IPR000560; -			
DR Pfam: PF00328; acid_phosphat; 1.			
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.			
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.			
DR Hydrolase; Glycoprotein; Signal; Multigene family.			
KW SIGNAL	1	17	BY SIMILARITY.
FT CHAIN	18	467	ACID PHOSPHATASE PHO11.
FT ACT_SITE	75	75	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE	337	337	PROTON DONOR (BY SIMILARITY).
FT CARBOHYD	97	97	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	162	162	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	192	192	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	250	250	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	315	315	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	356	356	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	390	390	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	439	439	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	445	445	N-LINKED (GLCNAC. .) (POTENTIAL).

Sat Oct 27 15:25:43 2001

us-09-488-265-27.rsp

```
Query Match          15.2%; Score 353; DB 1; Length 467;
Best Local Similarity 26.8%; Pred. No. 9.3e-22;
Matches 117; Conservative 62; Mismatches 187; Indels 70; Gaps 16;

QY 21 GOYSPFFSLADESATSPDPKGRVTFVQVLSRHGARYPTSSKSKYSALIERIQK-NAT 79
DB 42 GGAGPYFFPGDYGISRLDPESCEMKQLMARHGERYPYTSKCATIMKTWYKLSNTRQ 101
QY 80 FKGKYAFLLK-TYNYTLGADD-----LTPF-GENQMVNSGKIFRYRYKALAR 123
DB 102 FNGSLFLNDDYEFFIRDDDDLEMETTFANSNDVNLNPNVTGEMDAKRARELAQYWFYE 161
QY 124 NIVPFF-VRASGSDRVIAAEAFIEGFQSAKLADPAHQASPVINVIPEGSGYNNLTDHGL 182
DB 162 NOTSPPIFAASSERVHDTAQFIDG-----LGDQFNISLQIVSEAMSAGA---NTLSAGN 213
QY 183 CTFADSTLGDDAEANTAVFAPPIRAKL-RALPGVNLTDDEVNLMDCPFIVARTSD 241
DB 214 ACPGWDEDANDDILDKYDTTLLDDIAKRLNKENKGLNLTSDANTLFAWCAYELNAR--- 270
QY 242 ATQLSPFCLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVG-FNELIARLTHSPVQD 298
DB 271 --GYSDVCDIFTEDELVRYSYGODLVSFQDGPYDMIRSVGANLFNATLKLKQSETQD 328
QY 299 HTSTNHTLSDNPFPLNATLYADFSDHNTMYSIFFALGLYNGTKPLSTVSIESIETDGY 358
DB 329 -----LKVWLSFTHDITLNYLTAGIIDDKNLTAEYVPPFM-GNTF 369
QY 359 AASWTVPFAARAIVEMQCEAGGGGEGEKEPLRVLVNDRVPLVGHGCVDKLGRCKLDD 418
DB 370 HKSWMYPQGARVYTFEQC-----SNDTVYRVINDAVVPIETCTSGPGFSCIND 420
QY 419 F-----VEGLSFAR 427
DB 421 FDYAEKRVAGTDFLK 436

RESULT 7
ID PA5_YEAST STANDARD; PRT; 467 AA.
AC P00635;
DT 21-JUL-1986 (Rel. 01; Created)
DT 01-OCT-1994 (Rel. 30; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
GN PHO5 OR YBR093C OR YBR0814.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-26.
RX MEDLINE=83168913; PubMed=6300772;
RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
of repressible acid phosphatase contains a signal peptide."
RL Nucleic Acids Res. 11:1657-1672(1983).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-45.
RX MEDLINE=85037940; PubMed=6093051;
RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
RT "Structural analysis of the two tandemly repeated acid phosphatase
genes in yeast."
RL Nucleic Acids Res. 12:7721-7739(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA Mannheim G., Stucka R., Ehle S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II."
RL Yeast 10:1363-1381(1994).
RN [4]

SEQUENCE OF 1-44 FROM N.A.
MEDLINE=87064474; PubMed=3537710;
RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
RA Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
cluster within the acid phosphatase multigene family of Saccharomyces
cerevisiae."
Mol. Cell. Biol. 6:1855-1865(1986).
CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
ALCOHOL + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
CC -1- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; V01320; CAA24630.1; -
EMBL; X01079; CAA25555.1; -
EMBL; X78993; CAA55598.1; -
EMBL; Z35962; CAA85046.1; -
EMBL; X01080; CAA25556.1; -
PIR; S05795; PABYC.
PIR; S25241; B25241.
HSP; P34752; IHP.
SGD; S0000297; PHO5.
InterPro; IPR000560; -
Pfam; PF00328; acid_phosphat; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1
FT CHAIN 17
FT ACT_SITE 75
FT ACT_SITE 337
FT CARBOHYD 97
FT CARBOHYD 103
FT CARBOHYD 162
FT CARBOHYD 192
FT CARBOHYD 250
FT CARBOHYD 315
FT CARBOHYD 356
FT CARBOHYD 390
FT CARBOHYD 439
FT CARBOHYD 445
FT CARBOHYD 456
FT CARBOHYD 461
FT CONFLICT 36
FT CONFLICT 130
FT CONFLICT 294
FT CONFLICT 446
FT CONFLICT 462
FT CONFLICT 466
FT CONFLICT 467
SQ
REPRESSIBLE ACID PHOSPHATASE.
NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
PROTON DONOR (BY SIMILARITY).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
D -> Y (IN REF. 1).
A -> G (IN REF. 1).
H -> Q (IN REF. 1).
S -> V (IN REF. 1).
AS -> DT (IN REF. 1).
R -> K (IN REF. 1).
DC3C9504BC2D3D0C CRC64;

Query Match          15.0%; Score 347; DB 1; Length 467;
Best Local Similarity 26.5%; Pred. No. 2.9e-21;
Matches 117; Conservative 57; Mismatches 187; Indels 80; Gaps 18;

QY 21 GOYSPFFSLADESATSPDPKGRVTFVQVLSRHGARYPTSSKSKYSALIERIQK-NAT 79
DB 42 GGAGPYFFPGDYGISRLDPESCEMKQLMARHGERYPYTSKCATIMKTWYKLSNTRQ 101
QY 80 FKGKYAFLLK-TYNYTLGADD-----LTPF-GENQMVNSGKIFRYRYKALAR 123
DB 102 FNGSLFLNDDYEFFIRDDDDLEMETTFANSDDVNLNPNVTGEMDAKRARELAQYWFYE 161
```


QY 124 NIVPF-VRASGSDRVIAAEKFIIEGFSQAKLADPAHQASPVINVLII-----PEGSGYNNT 177
 Db 162 NOTSFVFTSNKRCHDTAQVFDG-----LQDQ-----FNITLTVSEAEAGANTL 209
 QY 178 LDHGLCTAFEDSLGDAAEFANFTAVFAPPIRARI-EALPGVNLTDVAVNLMDCPPDVT 236
 Db 210 SACNSCPAW-DYDANDDIVNEYDTTLLDDITAKRLNKENKGLNLTSTDAFLSWCAFEV- 267
 QY 237 ARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVG-FNELIARLTH 293
 Db 268 ----NAKGYSDVCDITKDELVHYSYQDLHTYHEGPGYDIKSGVSNLNFNASVKLLKQ 323
 QY 294 SPVQDHTSNHTLDSNPATFPLNATLYADFSDHTNTPSIFPAGLYNGTKPLSTTTSVESI 353
 Db 324 SEIQDQ-----KWLSTHTDITLTLTTAGIIDDKNNLTAEYVPM 365
 QY 354 ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGGVQDKLGR 413
 Db 366 -GNTFHRSMVVPQGARVYTEKFC-----SNDTYVRYVINDAVVPIETCTGTGFGFS 415
 QY 414 CKLDDE-----VEGLSPAR 427
 Db 416 CEINDFYAERKRVAGTDFLK 436

RESULT 8
 ID PPAD_YEAST STANDARD; PRT; 468 AA.
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE PROBABLE ACID PHOSPHATASE YDL024C (Last annotation update)
 GN YDL024C OR D2815.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B., Vissers S., Urrestarazu L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

DR EMBL; Z48432; CAA8335.1; -
 DR EMBL; Z74072; CAA98583.1; -
 DR SGD; S0002182; YDL024C.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multitene family.
 FT SIGNAL 1 20
 FT CHAIN 21 468
 FT ACT_SITE 76 76
 FT ACT_SITE 338 338
 FT CARBOHYD 98 98
 FT CARBOHYD 163 163
 FT CARBOHYD 193 193
 FT CARBOHYD 202 202
 FT CARBOHYD 238 238

FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDDF162C CRC64;

Query Match 14.5%; Score 337; DB 1; Length 468;
 Best Local Similarity 26.0%; Pred. No. 2e-20;
 Matches 119; Conservative 60; Mismatches 161; Indels 118; Gaps 19;

QY 21 GQVSPFSLADESAISPDVPGKGRVTFVQVLSRHGARYPTSSSKKYSALIERLOKNTWF 80
 Db 43 GGSAPYFSPANYGIPDTIDPEGCRLTQVMIGRHGERTYRSEAKD---IFEVYKISNY 99
 QY 81 KGRY----AFLKT-YNYTL-----GADDLTFP-GENQMVNSGKIFYRRIKA 120
 Db 100 TGYEGSLSFLNNGYEFFIPDESLEMETTLQNSIDLNPYTGEMNAKRHAREFLAKYK 159
 QY 121 LARNIVPF-VRASGSDRVIAAEKFIIEGFSQAKLADPAHQASPVINVLPEGSGYNNTLD 179
 Db 160 LMENCTNPFITNSKRIYDTAQYFAEL-----GDGFNISLQ 197
 QY 180 HGLCTAFEDSTLG-----DDAEANFTAVFAPPIRARI-EALPGVNL 220
 Db 198 ----TLSENSSGANTLAAKSCPNWNSNANDILMSYSDYLENLSDRNDENKGLNLS 253
 QY 221 DEDVYNLMDCPDTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGP 278
 Db 254 RKDAAALFSWCAFE-----NAKGYSNICDIFSAEALIHYSYETDLTSTFYQNGPGYKLIK 308
 QY 279 AQGVG-FN---ELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHTNMTVSIF 334
 Db 309 SIGANLEFNATVKLIQSAH-----LDQVWLSFTHDITDILNLT 347
 QY 335 ALGLYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRV 394
 Db 348 TAGLIDTRNLTTNIV-PFRHSVHRSWYIPOGARVYTEKFC-----SNDSYVRY 397
 QY 395 LVNDRVPLHGGVQDKLGRCKLDDE-----GLSF 425
 Db 398 VVNDVAVPIESCSGPGFSGEGTFYAYKDLRGVSF 435

RESULT 9

PPAL_PICPA STANDARD; PRT; 468 AA.
 ID PPAL_PICPA
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ACID PHOSPHATASE PHOI PRECURSOR (EC 3.1.3.2).
 GN PHOI.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E., Gannon P.M., Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RT characterization of the gene and its product.";
 RL Gene 163:19-26(1995).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: BY PHOSPHATE STARVATION.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Aspergillus awamori.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=105351;
[1]_

STRALINE=AKR0443;
 MEDLINE=94040796; PubMed=8224894,
 Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
 Middleton-Olmonen A., Nevalainen H., Ramosek J.A.,
 "The cloning and sequencing of the genes encoding phytase (phy) and
 pH 2.5-optimum acid phosphatase (aph) from *Aspergillus niger* var.
awamori,"
 Gene 133:55-62(1993).
 [2]
 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 MEDLINE=99264417; PubMed=10329192;
 Kostrewa D., Wyss M., D'Arcy A., van Loon A.P.;
 "Crystal structure of *Aspergillus niger* pH 2.5 acid phosphatase at
 2.4-A resolution,"
 J. Mol. Biol. 288:965-974(1999).
 -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYATE.
 -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.

-|- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the EMBL Outstation - Bioinformatics and the EMBL Outstation - Protein Crystallography.

```

-----
use the unprocessed data from the above mentioned institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/) or send an email to license@isb-sib.ch.
-----

EMBL: L02420; AAA16897.1;
PIR: JN0890; JN0890.
PDB: 1OFX; 19-APR-00.
InterPro: IPR000560;
Pfam: PF00328; acid_phosphat_1.
PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
Hydrolase; Glycoprotein; Signal; 3D-structure.
BY SIMILARITY.
SIGNAL 1 19
CHAIN 20 479
ACT_SITE 82 82
ACT_SITE 337 337
DISULFID 71 387
DISULFID 128 472
DISULFID 216 441
DISULFID 225 298
DISULFID 413 421
CARBOHYD 191 191
CARBOHYD 315 315
CARBOHYD 458 479
SEQUENCE 479 AA; 52678 MW; 4F8E0F3778CC3B08 CRC64;
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
3-PHYTASE B.
NUCLEOPHILIC ACCEPTOR.
PROTON DONOR.

```

Category	Score	DB 1	Length
Every Match	13.2%	307.5	479
Best Local Similarity	27.4%	Pred. No. 5.7e-18;	
Conservative	55	Mismatches 175;	Indels 99;
Indels	124	Gaps	23;

```

41 DGYSI--LKHVGNGPYSERVSY-----GIARDPTCEVDQVIMVKRGHVPSPSAGKD 94
67 YSALIETRIQK-NAT-FKGKVFALKTNYTL-----GADDLT-PF-GENQMVNSGKFFYR 116
95 IEEALAKVYSINTTEYKGGDLAFELNDWTYYVPNECYNAETSTQPYAGLGDAYNHGNDYKA 154
117 RYKAL--ARNIVPVRVRSAGSDRVITASAEKTEIEGFQAKLADPAHQASPVINVIPEGSGY 174

```

Db 155 RYGLHNGETVVPFF--SSGVRVETARKFGEF-----FGY 190

QY 175 NNTLDHGLCTAFEDSTLGDAAE-----NFTAVFAPPPIR---ARLEAL-PGV 217

Db 191 NYSNAALNIISEVGAUSLPTCTDNDQTTCDNLTYQL-PQFVAAARLNSQNGM 249

QY 218 NLTDDEVNLMDCPFTVARTSDATQSPFCDLFTADEW-QYDYLQSLKYGYGAG--- 273

Db 250 NLTASDVNLMVASFELNAR-----PFSNWINAFTQDEWVSFGYEDLNYI-YCAGPGD 303

QY 274 NPLGPAQGVGFENELIARLTHSPVQDHTSTNHTLDSNPATPPLNATLYADFSHNTMVSIF 333

Db 304 KMAAAGVAYANASITLLNQGPE-----AGSLFFNFAHDTNITPIL 345

QY 334 FALGLY--NGTKPLSTTSVESIETDGYAASVTPFAARAVVENMOCFAGGGGEGEKEPL 391

Db 346 AALGVLIPLNEDLPDRVAF-----GNPYSIGNIVPMGGHLTIERLSQATALSDEG---TY 398

QY 392 VRLVNDRVVPLHGGCVDKLGRCKLDDFVEGLS 424

Db 399 VRLVNEAVLPFNDCTSGPGYSCPLANYTSILN 431

RESULT 11

PHYB.ASPNG

ID PHYB.ASPNG STANDARD; PRT; 479 AA.

AC P34754;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE

DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE

GN PHYB.

OS Aspergillus niger.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI_taxid=5061;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.

RX MEDLINE=93371452; PubMed=7916610;

RA Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,

RA Ullah A.H.J.

RT "Identification and cloning of a second phytase gene (phyB) from

RL Biochem. Biophys. Res. Commun. 195:53-57(1993).

CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE

CC FROM PHYTATE.

CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =

CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.

CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L20567; AAA02934.1; -

CC InterPro: IPR000560; -

DR Pfam: PF00328; acid_phosphat; 1.

DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.

DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.

KW Hydrolase; Glycoprotein; Signal.

FT SIGNAL 1 19

FT CHAIN 20 479

FT ACT_SITE 81 81

FT REQUIRED FOR BINDING SUBSTRATE

FT (BY SIMILARITY).

FT ACT_SITE 82 82

FT ACT_SITE 337 337

FT ACT_SITE 106 106

FT PROTON DONOR (BY SIMILARITY).

FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 479 AA; 52611 MW; 395D4DA2B50DFC4 CRC64;

Query Match 13.2%; Score 305.5; DB 1; Length 479;

Best Local Similarity 27.4%; Pred. No. 8.4e-18;

Matches 124; Conservative 57; Mismatches 173; Indels 99; Gaps 24;

QY 9 DGYQCEPISHUWQ--YSPFSLADESAISPDYKCRVTFVQVLSRHGARYTSSKSK 66

Db 41 DGYSI--LKHGNGGYPYSERVSY---GIADPPTCEVDQVIMVKRIGERYSPSAGKS 94

QY 67 YSALIERIOK-NAT-PKGYAFILKTYNTL-----GADDLT-PF-GENOMVNSGIKEYR 116

Db 95 IEALAKVYSINTTEYKGLAFLNDWTYYVPNECYNAETTSQPYAGLLDAYNHGNDYKA 154

QY 117 RYKAL--ARNIVPEVRASGSDRVIASAEKIEGFQSAKLADPAHQASPVINVIPEGSGY 174

Db 155 RYGLHNGETVVPFF--SSGVRVETARKFGEF-----FGY 190

QY 175 NNTLDHGLCTAFEDSTLGDAAE-----NFTAVFAPPPIR---ARLEAL-PGV 217

Db 191 NYSNAALNIISEVGAUSLPTCTDNDQTTCDNLTYQL-PQFVAAARLNSQNGM 249

QY 218 NLTDDEVNLMDCPFTVARTSDATQSPFCDLFTADEW-QYDYLQSLKYGYGAG--- 273

Db 250 NLTASDVNLMVASFELNAR-----PFSNWINAFTQDEWVSFGYEDLNYI-YCAGPGD 303

QY 274 NPLGPAQGVGFENELIARLTHSPVQDHTSTNHTLDSNPATPPLNATLYADFSHNTMVSIF 333

Db 304 KMAAAGVAYANASITLLNQGPE-----AGSLFFNFAHDTNITPIL 345

QY 334 FALGLY--NGTKPLSTTSVESIETDGYAASVTPFAARAVVENMOCFAGGGGEGEKEPL 391

Db 346 AALGVLIPLNEDLPDRVAF-----GNPYSIGNIVPMGGHLTIERLSQATALSDEG---TY 398

QY 392 VRLVNDRVVPLHGGCVDKLGRCKLDDFVEGLS 424

Db 399 VRLVNEAVLPFNDCTSGPGYSCPLANYTSILN 431

RESULT 12

PPA5_KLULA

ID PPA5_KLULA STANDARD; PRT; 469 AA.

AC P52289;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).

GN PHO5.

OS Kluyveromyces lactis (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.

OX NCBI_taxid=28985;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CBS 2359/152;

RA Ferminan E.

RL Thesis (1995), University of Salamanca, Spain.

CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN

CC ALCOHOL + ORTHOPHOSPHATE.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).

CC -1- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.

CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 233995; CA83964.1; --
InterPro; IPR000560; --
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 16
FT CHAIN 17 469
FT ACT_SITE 77 77
FT ACT_SITE 339 339
FT CARBOHYD 23 23
FT CARBOHYD 31 31
FT CARBOHYD 32 32
FT CARBOHYD 129 129
FT CARBOHYD 201 201
FT CARBOHYD 229 229
FT CARBOHYD 250 250
FT CARBOHYD 317 317
FT CARBOHYD 392 392
FT CARBOHYD 447 447
FT SEQUENCE 469 AA; 52532 MW; 863B528D0740AA7E CRC64;

Query Match 12.9%; Score 300.5; DB 1; Length 469;
Best Local Similarity 25.9%; Pred. No. 2.1e-17;
Matches 118; Conservative 66; Mismatches 17; Indels 99; Gaps 21;

QY 17 SHLWGOYSPFFSLADESATSPDVPKGRVTFVQLSRHGARYPTSSKSKYSALIERIQK 76
DB 41 SLNGO-GPHYDPSQFGLPEVDPDCTVEHOMLARHGERYPTASKGKLMIALMDKLKE 99
QY 77 NATFKGYAFKTYNTLGGADLTTPGEMQVNSGKIFRYR 118
DB 100 ---FOGOY-----NDPLEVFNDEFFVSNTRYFDQLTNSTVDPSNPYAGAKTA 145
QY 119 ----KALARNIVPVFRASGSDRVIAAEKFIQFQSAKLADPAHQASPVINV---IIPEG 171
DB 146 QHLGKIYAYNGDLF--SDSNPVFTSSGRVH--QTAKYVWSLEELDIQLDLQIQEN 201
QY 172 --SGYNTLDHGLCTAFEDSTLGDGAENFTAVFPIRAR-LEALPGVNL--DEDVYN 226
DB 202 ETSANSLTPADSCWTY-NGDLGDEVFENATLPLTDIKNRNMKNNSNLTLEHDDIEL 260
QY 227 LMDMCPDVTARTSDATQLSPFCDLFTADEW-OYDYLOSL-KYGYGAGNPL-GPAQGVG 283
DB 261 LVDWCAFEETNVKGSAY-----CDLFRNDLVAYSYVANNFYRRGAGNPMSPIGSVL 315
QY 284 FNELIARLTHSPVDHSTNHTLDSNPATFPLNATLYADFSDHNTWISVIFALGLY-NGT 342
DB 316 VNASYNLLTQADELN-----KWLSESHDITQOQFISALGLIDNGV 357
QY 343 KPLSTTSV--ESITDGYASWTVPPFAARAYVENMQCEAGGGEKEPLVRVLVNDRV 400
DB 358 TEYSLDQVDFQNIQ-----QLSWVTPMGRIFTKLKC-----GNASVRYIINDVI 404
QY 401 VPLHCGGVGDKLGRCKLDDF-----VEGLSFARS 428
DB 405 IPVPGCTSGPFCSEPIEDFDYIINRLNGIDYVSS 439

RESULT 13
PPAL SCHPO STANDARD; PRT; 453 AA.
ID PPAL SCHPO
AC P08091;
DT 01-AUG-1988 (rel. 08, Created)
DT 01-AUG-1988 (rel. 08, Last sequence update)

DT 01-OCT-2000 (rel. 40, Last annotation update)
DE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
GN PHO1 OR SPBP4G3.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86140050; PubMed=3005272;
RA Elliott S., Chang C., Schweingruber M.E., Schaller J., Rickli E.E.,
RA Carbon J.;
RT "Isolation and characterization of the structural gene for secreted acid phosphatase from Schizosaccharomyces pombe";
RL J. Biol. Chem. 261:2936-2941(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RI Rieger M., McDougall R.C., Rajandream M.A., Barrell B.G.;
RI Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: CELL WALL.
CC -!- INDUCTION: RERESSED BY PHOSPHATE AND WEAKLY BY THIAMINE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M11857; AAA35321.1; --
EMBL; AL137099; CAB68657.1; --
PIR; A25326; A25326.
DR HSP; P34752; IHP.
DR InterPro; IPR000560; --
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Cell wall; Signal.
FT SIGNAL 1 18
FT CHAIN 19 453
FT ACT_SITE 69 69
FT ACT_SITE 330 330
FT CARBOHYD 95 95
FT CARBOHYD 151 151
FT CARBOHYD 183 183
FT CARBOHYD 193 193
FT CARBOHYD 243 243
FT CARBOHYD 319 319
FT CARBOHYD 410 410
FT CARBOHYD 429 429
FT CARBOHYD 443 443
FT SEQUENCE 453 AA; 50557 MW; 7CF891256EB154D1 CRC64;
Query Match 12.8%; Score 298; DB 1; Length 453;
Best Local Similarity 25.6%; Pred. No. 3.2e-17;
Matches 107; Conservative 63; Mismatches 182; Indels 66; Gaps 16;

QY 23 YSPFFSLADESATSPDVPKGRVTFVQLSRHGARYPTSSKSKYSALIERIQK---NAT 79
DB 43 HKPYF-----YGPSTDFPTTKIKQVHTLQRHGRNPTGNGNAFVAVGLANFQORLLNGS 97
QY 80 FKGY-----AFLKTYNTL---GADLTTFGENQVNSGKIFRYRYKALARNIVPV 129
DB 98 VPTIDYVSGNPLSFVPTWTPTVIEANADALSSGRVLEFDMGRQFYRYHELNFNASTYNI 157
QY 130 RASGSDRVIAAEKFIQFQSAKLADPAHQASPVINVIPEGSYNTLDHGLCTAFEDS 189

Db 158 YTAQQRVDSALWYGYCM-----FGEDVHNFTNVLISENATAGSNLSNACPASDAD 213
 QY 190 TLGDDAEANFTAVAPTRARLEA-LPGVNLTDDEVNLMDCMDFIVARTSDATQLSPF 248
 Db 214 DFTTAPALEARNVTPPIKRLNPNDFNSNLTNDIILNLYGICSYETALQ-----DYSEF 268
 QY 249 CDLFTADWQYDYLQSLKYGYGAGNPL--GPAQGVGFENELIARLTHSPVQDHTSNHT 305
 Db 269 CKLFNSVDLNFVEGDLG--FSYGMGNSVKGWGSIFGAYANSIANSURS-VENNTO----- 322
 QY 306 LDSNPATFPLNATLYADFSHDNTMWSIFFALGYNGT---KPLSTTSVESIETDGYAASW 362
 Db 323 -----QVFTAFTHDANIIPVETALGFDTONTPENPLPTS--YQVHSHMKASE 368
 QY 363 TVPFAARAYVEMMQCAGGGGCGEKEPLVRVLVNDRVVPLHCGGVYDKL-----GRCKL 416
 Db 369 FVPFAGNLITELPOCE-----DSKYVVRHLVNEEVPLSDCGFGPSNTSDGMCEL 418

RESULT 14
 ID PPA2_SCHPO STANDARD; PRT; 463 AA.
 AC Q01682;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 GN PH04 OR SPBC428.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
 RA MEDLINE=91064763; PubMed=2249257;
 RX Yang J., Schweingruber M.E.; for thiamin-repressible acid phosphatase
 RT "The structural gene coding for thiamin-repressible acid phosphatase
 RL in Schizosaccharomycetes pombe.";
 RN Curr. Genet. 18:269-272(1990).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- INDUCTION: REPRESSED BY THIAMINE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56939; CAA0258.1;
 DR EMBL; AL034382; CAA22278.1;
 DR PIR; S14119; S14119.
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 463
 FT ACT_SITE 69 69
 FT ACT_SITE 340 340
 FT THIAMINE-REPRESSIBLE ACID PHOSPHATASE.
 FT NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT PROTON DONOR (BY SIMILARITY).

FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 463 AA; 52118 MW; F48EAPF8B6B6234A CRC64;

Query Match 12.5%; Score 290; DB 1; Length 463;
 Best Local Similarity 24.3%; Pred. No. 1.5e-16;
 Matches 109; Conservative 70; Mismatches 183; Indels 86; Gaps 18;

QY 23 YSPFESLADSSALSPDVPKCRVTFVQVLSRHGARYPTSS-----KSKYKYSALIERIQK 76
 Db 43 HEPIYF-----NGPTTFPESCAIKQVHLLQHGSRNPTGDDTATDVSSAQYIDIFQNKLL 97
 QY 77 NATFFKGKYA-----FLKTYNYTL--GADDLTPGENOMVNSGKIFYRRYKALARNIV 126
 Db 98 NGSIPTVNESYPENPLYFVKHWTPIKAENADQLSSGRIELFDLGRQVFFERYYELEDTV 157
 QY 127 PFVRASGSDRVIAAEKFTIEGFOSAKLADPAHQASPVINVIPE--GSGYNNTLDHGLCT 184
 Db 158 YDINTAAQERVVDSAEWFSYMGFDDMQNKTN-----FIVLPEDDSAGANSAMYYSCP 211
 QY 185 AFEDSTLGGD---AEANFTAVFAPPFIRARLEAL--PGVNLTDDEVNLMDCMFPFTVAR 238
 Db 212 VYEDNNIDENTTEAHTSWRNVLKPLIANLNKYFDSGYNLVTSVDVRSLYICVYETIALR 271
 QY 239 TSDATQLSPFCDLFTADEW--QYDYLQSLKYGYGAGNPLGPAO-----GVCFNELIAR 290
 Db 272 DN-----SDFCSLFTPSEPLNEFYDSDLDYVWG-----GPASEWASTLCGAYVNNLANN 321
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMWSIFFALGLY---NGTKPLST 347
 Db 322 L-----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDITPEHPPLT 365
 QY 348 TSVESIETDGYAASWTVPFAARAYVEMMQCAGGGGGEKEPLVRVLVNDRVVPLHCGC 407
 Db 366 D--KNIFTYSLKTSFSFVPFAGNLITELFLC-----SDNKYYVVRHLVNOQVYPLTDCG 415
 QY 408 VDKLGR-----CKLDDFVEGLSFAFSGGN 431
 Db 416 YGPGASDGLCELSELYNSVVRVNSTN 443

RESULT 15
 ID PPAX_CAEEL STANDARD; PRT; 755 AA.
 AC Q09549; O17843;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE ACID PHOSPHATASE F26C11.1 (EC 3.1.3.2).
 GN F26C11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RC Matthews P., Lloyd C.;
 RA Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:24 ; Search time 75.85 Seconds
(without alignments)
762.258 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDTVDGQCPEISHLW.....DFVEGLSPARSGGNWAECA 437

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SP TREMBL_16.*
- 2: SP archaea.*
- 3: SP bacteria.*
- 4: SP fungi.*
- 5: SP human.*
- 6: SP invertebrate.*
- 7: SP mammal.*
- 8: SP mhc.*
- 9: SP organelle.*
- 10: SP plant.*
- 11: SP rodent.*
- 12: SP unclassified.*
- 13: SP vertebrate.*
- 14: SP virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1762.5	75.9	465	3	O00092
2	1683	72.5	467	3	Q9U027
3	1672	72.0	467	3	Q93838
4	1662	71.6	467	3	Q9HQ00
5	1596.5	68.8	466	3	O00096
6	1588	68.4	466	3	O00100
7	1562.5	67.3	466	3	O00085
8	1275.5	55.0	487	3	O00107
9	317.5	13.7	442	3	Q74677
10	298.5	12.9	469	3	Q9W846
11	285	12.3	463	3	O60172
12	207.5	8.9	460	5	Q9VV72
13	196	8.4	467	5	O96421
14	192.5	8.3	453	5	O96420
15	191.5	8.3	453	5	Q9W438
16	176	7.6	487	4	Q9UGA3
17	175	7.5	451	11	O35217
18	174	7.5	487	4	Q9UNW1
19	173	7.5	487	4	O95172

20	172	7.4	481	11	Q9Z2L6
21	166	7.2	449	13	Q92170
22	154	6.6	468	10	O04509
23	142	6.1	198	3	Q9UTX1
24	133	5.7	274	11	Q9JUD5
25	129.5	5.6	381	11	Q9JGD5
26	128.5	5.5	374	11	Q9JMG5
27	115	5.0	683	5	O00838
28	113	5.0	707	5	O00839
29	111	4.8	452	5	Q19175
30	109.5	4.7	447	5	Q9US59
31	106.5	4.6	447	5	Q9US75
32	106.5	4.6	447	5	Q9US58
33	106.5	4.6	447	5	Q9TW19
34	106.5	4.6	447	5	Q9TW17
35	106	4.6	2205	5	O16857
36	105.5	4.5	447	5	Q97187
37	105.5	4.5	447	5	Q9U5T4
38	105.5	4.5	447	5	Q9U5T0
39	105.5	4.5	447	5	Q9TW40
40	105	4.5	537	5	O25332
41	105	4.5	888	5	O25336
42	104.5	4.5	447	5	Q9U5V1
43	104.5	4.5	447	5	Q9U5U7
44	104.5	4.5	447	5	Q9U5U3
45	104.5	4.5	447	5	Q9U5T3

ALIGNMENTS

RESULT 1
O00092
ID O00092 PRELIMINARY; PRT; 465 AA.
AC O00092;
DT 01-JUL-1997 (TREMREL. 04, Created)
DT 01-JUL-1997 (TREMREL. 04, Last sequence update)
DT 01-MAY-2000 (TREMREL. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (BC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHVA.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_taxid=5085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN=ATCC 34625;
RX MEDLINE=97288063; PubMed=9143104;
RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
RT phytase from the fungus Aspergillus fumigatus.";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).
CC -|- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -|- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -|- SUBCELLULAR LOCATION: SECRETED.
CC -|- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
CC A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
CC 7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMIC ACTIVITY.
CC -|- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; U59804; AAB96872.1; -;
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 26

Q9Z2L6 mus musculus
Q92170 gallus gall
O04509 arabidopsis
Q9UTX1 schizosacch
Q9JJD5 mus musculus
Q9JGD5 mus musculus
Q9JMG5 mus musculus
O00838 leishmania
O00839 leishmania
Q19175 caenorhabdi
Q9US59 drosophila
Q9US75 drosophila
Q9US58 drosophila
Q9TW19 drosophila
Q9TW17 drosophila
O16857 microciona
Q97187 drosophila
Q9U5T4 drosophila
Q9U5T0 drosophila
Q9TW40 drosophila
Q25332 leishmania
Q25336 leishmania
Q9U5V1 drosophila
Q9U5U7 drosophila
Q9U5U3 drosophila
Q9U5T3 drosophila

Sat Oct 27 15:25:44 2001

```

FT CHAIN 27 465 3-PHYTASE A. 1762.5; DB 3; Length 465;
FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
FT SIMILARITY).
FT 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT 359 359 PROFON DONOR (BY SIMILARITY).
FT 39 BY SIMILARITY.
FT 70 70 BY SIMILARITY.
FT 412 412 BY SIMILARITY.
FT 463 463 BY SIMILARITY.
FT 262 262 BY SIMILARITY.
FT 434 434 BY SIMILARITY.
FT 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 465 AA; 50836 MW; 86FCID9058C9B2C9 CRC64;

Query Match 75.9%; Score 1762.5; DB 3; Length 465;
Best Local Similarity 77.4%; Pred. No. 7.7e-136;
Matches 345; Conservative 36; Mismatches 48; Indels 17; Gaps 10;

QY 2 SHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQLSRHGARYP 59
DB 27 SKSCDVTDLGYQCSPATSHLWGOYSPFFSLADESAISPDVPGKCRITLTVQLSRHGARYP 86
QY 60 TSSKSKKYSALIERIOKNA-TEFGKYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 118
DB 87 TSSKSKKYSALIERIOKNA-TEFGKYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 146
QY 119 KALARNIVPVRASGSDRVIASAEKFIIEGFSQAKLADP--AHQASPVINVIIPESGYNN 176
DB 147 KALARSVVPFIRASGSDRVIASAEKFIIEGFSQAKLADPAGATRAAPAIIVIPESGFENN 206
QY 177 TLHGLCTAFEDSTLGDAAEFANFTAFAPFARLEA-LPGVNLTDDEVVNLMDMCPFT 235
DB 207 TLHGVCTFEASQOLGDEVAANTFALFADIRAEKHLPGVTLTDEDVSLMDMCSFT 266
QY 236 VARTSDATQLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVF-NELIARLT 292
DB 267 VARTSDASQLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVF-NELIARLT 326
QY 293 HSPVQDHTSTNHLDSNPATFPLNATLYADFSDNTMVSIFLALGLYNGTKPLSTTSV 352
DB 327 RSPVQDHTSTNHLDSNPATFPLNATLYADFSDNTMVSIFLALGLYNGTKPLSTTSV 386
QY 353 I-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVDKL 411
DB 387 AKELDGYASWVFPFARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVDKL 439
QY 412 GRCKLDDEVEGLSFARSGGNWAECPA 437
DB 440 GRCKLDNDFVKGLSWARSGGNWAECPA 465

```

```

RESULT 2
Q90U027 PRELIMINARY; PRT; 467 AA.
AC Q90U027
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (EC 3.1.3.8).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RA Honning W., Qi W., Jing X.;
RT "PCR, cloning and characterization of the phytase (phyA) gene of

```

```

RT Aspergillus niger (China Strain).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218813; AAF25481.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000360; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Signal; Lyase; Hydrolase.
FT SIGNAL 1 19 POTENTIAL.
FT SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 72.5%; Score 1683; DB 3; Length 467;
Best Local Similarity 73.7%; Pred. No. 2.4e-129;
Matches 330; Conservative 43; Mismatches 57; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQLSRHGARY 58
DB 27 NQSTCDTVDQGYQCFSETSHLWGOYAPFFSLANESAISSPDVPGACRVTFVQLSRHGARY 86
QY 59 PTSSKSKKYSALIERIOKNA-TEFGKYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 117
DB 87 PTDSKSKKYSALIERIOKNA-TEFGKYAFILKTYNYTLGADDLTPFGENOMVNSGKIFRYR 146
QY 118 YKALARNIVPVRASGSDRVIASAEKFIIEGFSQAKLADP--AHQASPVINVIIPESGY 174
DB 147 YESLTRNIIPIRSSGSSRVIASGKFIIEGFSQAKLADPRAQPGQSPKIDVVISEASS 206
QY 175 NNTDLGLCTAFEDSTLGDAAEFANFTAFAPFARLEA-LPGVNLTDDEVVNLMDMCP 233
DB 207 NNTDLPGCTCTVFESLADAEVAFANFTATFTPTIRQRLENLDSVSLTDTETVYLMDCSF 266
QY 234 DTVARTSDATQLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVF-NELIAR 290
DB 267 DTISTSTVDTKLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVF-NELIAR 326
QY 291 LTHSPVQDHTSTNHLDSNPATFPLNATLYADFSDNTMVSIFLALGLYNGTKPLSTTSV 350
DB 327 LTHSPVQDHTSTNHLDSNPATFPLNATLYADFSDNTMVSIFLALGLYNGTKPLSTTSV 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVVPVPLHGCGVD 409
DB 387 QNTITDGFSSANTVPFASRLYVEMMOCEA-----EQEPLRVLVNDRVVPVPLHGCPVD 439
QY 410 KLGRCKLDDEVEGLSFARSGGNWAECPA 437
DB 440 ALGRCTRDTSFVKGLSFARSGGNWAECPA 467

```

```

RESULT 3
O93838 PRELIMINARY; PRT; 467 AA.
ID O93838
AC O93838;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE PHYTASE.
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eukaryotes; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK-57;
RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
RT "Phytase having high-affinity for phytic acid.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022700; CAB19824.1; -.
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.

```


Downloaded from ascelibrary.org by University of California, San Diego on 06/01/15. Copyright ASCE, For All Rights Reserved, No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage or retrieval system, without permission in writing from ASCE.

1

K	interPro; IPR000560; -
R	Pfam; PF00328; acid_phosphat: 1
R	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
R	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
W	Hydrolase; Glycoprotein; Signal.
S	SIGNAL
T	CHAIN
T	1
T	14
T	POTENTIAL
T	15
T	466
T	CHAIN
T	79
T	ACT_SITE
T	357
T	ACT_SITE
T	28
T	DISULFID
T	68
T	DISULFID
T	212
T	DISULFID
T	461
T	DISULFID
T	261
T	278
T	BY SIMILARITY.
T	BY SIMILARITY.
T	BY SIMILARITY.
T	BY SIMILARITY.
T	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
T	PROTON DONOR (BY SIMILARITY).
T	BY SIMILARITY.
T	BY SIMILARITY.
T	BY SIMILARITY.
T	BY SIMILARITY.

Sat Oct 27 15:25:44 2001

```

FT DISULFID 432 440 BY SIMILARITY.
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 AA; 51450 MW; FC4575B521A5C929 CRC64;
SQ SEQUENCE 466 AA; 51450 MW; FC4575B521A5C929 CRC64;

Query Match 68.8%; Score 1596.5; DB 3; Length 466;
Best Local Similarity 70.8%; Pred. No. 2.9e-122; Indels 19; Gaps 11;
Matches 317; Conservative 36; Mismatches 76;

QY 1 NSHSCDTVD-GYQC-PEISHLWGQYSPFSLADESAISPDPVPGKCRVTFVQVLSRHGARY 58
DB 24 DSHSCNTVEGGQCRPEISHSGWGQYSPFSLADQSEISPDVPQNCKITFVOLLSSRHGARY 83

QY 59 PTSSSKKYSALIERIQKNAT-FKGKIAFLKTYNTLGGADLTTPFGENQMVNSGKIFYRR 117
DB 84 PTSSSKKYSALIERIQKNAT-FKGKIAFLKTYNTLGGADLTTPFGENQMVNSGKIFYRR 143

QY 118 YKALARNIVPFVRASGSDRVIASAEKFIIEGFSQAKLADP---AHOASPVINVIIEGSGY 174
DB 144 YKALARNIVPFVRASGSDRVIASAEKFIIEGFSQAKLADP---AHOASPVINVIIEGSGY 203

QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFVAPPTRARL-BALPGVNLTDDEVDVNLMDMCPF 233
DB 204 NNTLDHGLCTAFEDSTLGDAAEANTFVAPPTRARL-BALPGVNLTDDEVDVNLMDMCPF 263

QY 234 DTVARTSDATQSPFCDLFTADEMQ-YDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 264 ETLLARNHTDT-LSPFCALSTQEEQAYDYQSLGKYGNGGNPLGPAQGVGFVNLIIAR 322

QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPFGALGXNGTKPLSTTSV 350
DB 323 MTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPFGALGXNGTKPLSTTSV 382

QY 351 EST-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGGCV 409
DB 383 KSIETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGGCV 435

QY 410 KLGCRKLDLDFVEGLSFARSGGNWAECEFA 437
DB 436 SLGRCKRDDDFVRLGLSEAROGNWECEYA 463

RESULT 6
ID O00100 PRELIMINARY; PRT; 466 AA.
AC O00100;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE).
OS Aspergillus terreus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=33178;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=CBS 116.46;
RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B.,
RA Broger C., van Leen A.P.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DB EMBL; U60412; AAB58465.1; -.

```

```

DR HSSP; P34752; 1IHP.
DR InterPro; IP000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 466 3-PHYTASE.
FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY
FT ACT_SITE 83 83 SIMILARITY).
FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
FT DISULFID 31 40 BY SIMILARITY.
FT DISULFID 71 414 BY SIMILARITY.
FT DISULFID 215 465 BY SIMILARITY.
FT DISULFID 264 282 BY SIMILARITY.
FT DISULFID 436 444 BY SIMILARITY.
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 466 AA; 51055 MW; F2ABCECIAF7C22C4 CRC64;

Query Match 68.4%; Score 1588; DB 3; Length 466;
Best Local Similarity 69.6%; Pred. No. 1.4e-121;
Matches 311; Conservative 46; Mismatches 72; Indels 18; Gaps 10;

QY 1 NSHSCDTVD-GYQC-PEISHLWGQYSPFSLADESAISPDPVPGKCRVTFVQVLSRHGARY 58
DB 27 NHDSDCTSDRGYQCPPELSHKWGLYAPFSLQDESPPFLDPDPDCHITFVQVLAHGAARS 86

QY 59 PTSSSKKYSALIERIQKNAT-FKGKIAFLKTYNTLGGADLTTPFGENQMVNSGKIFYRR 117
DB 87 PTDSKTAYATIAAIQKNATLPGKIAFLKSYNYSMSSENLPFGNQLDQLCAQFVRR 146

QY 118 YKALARNIVPFVRASGSDRVIASAEKFIIEGFSQAKLADP---AHOASPVINVIIEGSGY 174
DB 147 YDTLTHINPFVRAADSSRVHSAEKFVEGFQONARQDGFPHANPHQSPRVDVVIPEGTAY 206

QY 175 NNTLDHGLCTAFEDSTLGDAAEANTFVAPPTRARLEA-LPGVNLTDDEVDVNLMDMCPF 233
DB 207 NNTLEHSICTAFESTVGDAAADNFTAVFAPAIARLEADLPGVQLSADDDVVNLMMAMCPF 266

QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 267 ETVSLTDDAHTLSPFCDLFTAAEWQYNYLLSLDKYGYGGNPLGPAQGVGWANELIAR 326

QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPFGALGXNGTKPLSTTSV 350
DB 327 LTRSPVHDHTCVNNTLDANPATFPLNATLYADFSDNSNLVSIFWALGXNGTKPLSQTTV 386

QY 351 EST-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLRVLVNDRVPLHGGCV 409
DB 387 EDITRTDGYAANTVFAARAYIEMMQCRA-----EQPLRVLVNDRVPLHGGCAVD 439

QY 410 KLGCRKLDLDFVEGLSFARSGGNWAECEFA 436
DB 440 NLGRCKRDDDFVRLGLSEARAGGNWAECEFA 466

RESULT 7
ID O00085 PRELIMINARY; PRT; 466 AA.
AC O00085;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-

```

QY	235	TVARTSDATQLSPFCDLFTADW-QYDYLOSL-KYYGTGAGNPLGAQGVGF-NELIARL	291
Db	268	TVSLTDDAHTLSPPFCDLFTATEWTQNYLLSLDKYTYGGGNPLGPVGQWANELMARL	327
QY	292	THSPVDHSTNHNLDSNPATFPLNATLYADFSDNTMVSIFALGLNGTKPLSTTSVE	351
Db	328	TRAPVHDHTCVNNVLDSAPTEPFLNATLYADFSDNSLVSWALCLNGTAPLSQTISVE	387
QY	352	SI-ETDGYAASWTPEARAAYVMNQCEAGGGGEKEPELVRLVNDVRVPLHCGGVDK	410
Db	388	SYSOTDGYAAWTPEARAAYVMNQCR-----EXEPIRVRLVNDVRVPLHGCPDTK	440
QY	411	LGRCKLDFFVEGLSFARSGGNWAECF	436
Db	441	LGRCKRDFAVGLSFAQAAGNWADC	466

RESULT	8
O00107	
ID	O00107 PRELIMINARY; PRT; 487 AA.
AC	O00107;
DT	01-JUL-1997 (TREMBLrel. 04, Created)
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE	01-WAY-2000 (TREMBLrel. 13, Last annotation update)
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
DE	PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISP
DE	PHOSPHOHYDROLASE A).
GN	PHYA.
OS	Thielavia heterothallica.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC	Sordariales; Chaetomiaceae; Thielavia.
OX	NCBI_TaxID=78579;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97177792; PubMed=9025298;
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,
RA	van Loon A.P.G.M.;
RT	"The phytase subfamily of histidine acid phosphatases: isolation of
RT	genes for two novel phytases from the fungi Aspergillus terreus and
RT	Mycelophthora thermophila.";
RL	Microbiology 143:245-252(1997).
CC	-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC	FROM PHYTATE.
CC	-!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISP
CC	INOSITOL 1,2,4,5,6-PENTAKISP
CC	SUBCELLULAR LOCATION: SECRETED.
CC	-!- SUBCELLANEUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
CC	3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO
CC	ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA
CC	SHIFTED TO MORE ACIDIC PH VALUES.
CC	-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR	EMBL: U59806; AB52508.1; -
DR	HSSP: P34752; IHP.
DR	InterPro: IPR000560; -
DR	Fram; PF00328; acid_phosphat_1.
DR	PROSITE; PS00616; HIS_ACID_PHOSPAT_1; 1.
DR	PROSITE; PS00778; HIS_ACID_PHOSPAT_2; UNKNOWN_1.
KW	Hydrolase; Glycoprotein; Signal.
FT	SIGNAL 1
FT	CHAIN ? POTENTIAL.
FT	DOMAIN ? 487 3-PHYTASE A.
FT	POLY-SER.
FT	POLY-GLY.
FT	ACT_SITE 75 75 REQUIRED FOR BINDING SUBSTRATE (BY
FT	SIMILARITY).
FT	ACT_SITE 76 76 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE 368 368 PROTON DONOR (BY SIMILARITY).
FT	DISULFID 26 35 BY SIMILARITY.
FT	DISULFID 64 421 BY SIMILARITY.
FT	DISULFID 208 485 BY SIMILARITY.
FT	DISULFID 260 289 BY SIMILARITY.
FT	DISULFID 456 464 BY SIMILARITY.
FT	CARBOHYD 165 165 N-LINKED (GLCNAC...) (POTENTIAL).
FT	CARBOHYD 200 200 N-LINKED (GLCNAC...) (POTENTIAL).

Sat Oct 27 15:25:44 2001

```

Query Match      13.7%; Score 317.5; DB 3; Length 442;
Best Local Similarity 25.8%; Pred. No. 8.7e-18;
Matches 105; Conservative 67; Mismatches 158; Indels 77; Gaps 18;

QY 38 DVPKGRVTFVQVLSRHGARYPTSSKYSALIERIOK-NAT-FKGKAYELKTYNYTLG 95
Db 55 DTPHCEIEQAOLEFMRHGERPFTKSGQYKFKYDKLKKANITDYKGLAFIEDLEYFVP 114
QY 96 ADDLPFGQNMVNSGI---KF---YRRYKAL--ARNIVPFVRASGSDRVIASAEKFI 145
Db 115 DSDNYELETTRGLYSGLLNAFKFYRLRYRYSLVDTSSVLPFAAS-EDRVVDTFARFSG 173
QY 146 EGFSQAKLADPAHQASPVINVI---IPEGSYNNNTLDHGLC-----TAFEDSTLGDREA 197
Db 174 RGF-----FGPDYATSCSIQVNETDTSKGNALTTKDN--CPTYNSSFYDYSFGDE--- 223
QY 198 NETAFAPIRABLEALPGVNLTDVNVLMDCPDVTARTSDATQLSPPCDLFTADBEW 257
Db 224 ---IFQREADRLNELSPGNITADDIITMGTYCAYETNVKGH-----SSFCDALESREAF 274
QY 258 ---OYDYLSQSKYGYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPP 314
Db 275 IALQYNN-DVTKFYQFGPGYNMSAVAGVYANATAKLQE----- 313
QY 315 LNATLYADFSDHNTMVSIFPALLGYNGTKPLSTTSVESIETDGYAASWTVPFAARAYVEM 374
Db 314 -DGKLMFSFSDNDLLNYITALLGLITDTE-LGTEDVDHFRS---FKTSELVPOGARLIIEK 369
QY 375 MOCAGGSGGGEKEPELVRVLVNDVRVPLHCGGVKDLGRCKLDDFVE 421
Db 370 LNC-----SDTSFVRILNDKVPVPGCSSGPGYSCPLEDYLD 407

RESULT 10
QY846 PRELIMINARY; PRT; 469 AA.
ID QY846
AC QY846; 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ACID PHOSPHATASE (EC 3.1.3.2).
GN PHO3.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Kluyveromyces.
OC Saccharomycetaceae; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2360/7;
RA San Vicente A., Ferminan E., Dominguez A.;
RT "Isolation and characterization of KIPHO3 a gene encoding a
constitutive acid phosphatase from Kluyveromyces lactis."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ007502; CAB46490.1; -.
DR HSSP; P34755; IQFX.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid.phosphat. 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase.
FT CHAIN
SQ SEQUENCE 469 AA; 52560 MW; 5C7ABF622CEA891C CRC64;

Query Match      12.9%; Score 298.5; DB 3; Length 469;
Best Local Similarity 26.2%; Pred. No. 3.4e-16;
Matches 117; Conservative 65; Mismatches 174; Indels 91; Gaps 21;

QY 21 GOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTSSKYSALIERIOKNAIF 80
Db 45 GO-GPHYDYFQSGFIPVEVDQCTVEHVMARHGRIYPTASKGLWIALWDLKKE---F 100
QY 81 KGKY-----AFLKTYNYTLGADDLTP-----FGENQMVNSGIKIFRYRYKALA 122

Query Match      12.9%; Score 1275.5; DB 3; Length 487;
Best Local Similarity 56.3%; Pred. No. 5.2e-96;
Matches 263; Conservative 52; Mismatches 119; Indels 33; Gaps 13;

QY 2 SHSCDTPD-GYQC-PEISHLMGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYP 59
Db 23 SRPCTDPLDGCQTGTAISHFGQYSPFVSFSE--LDASIPDDCEVTFQAQVLSRHGARAP 80
QY 60 TSSKSKYSALIERIOKNAIFKFKG-YAFKLKTYNYTLGADDLTPFGENQMVNSGIKIFRYR 118
Db 81 TLKRAASVVDLIDRLHGAISYGPGEYELRYDYLGADELTRTGQOQMVNSGIKIFRYR 140
QY 119 KALARNIVPFVRASGSDRVIASAEKFIPEGFSQAKLADPAHQASPI---NVIIEGSGYN 175
Db 141 RALARKSIPFVRTAGODRVHSAENFTQGFHSALLADRGSTVRPTLPYDMVVIIPETAGAN 200
QY 176 NTLHGLCTAFED---STLGGDAEAFNTAVFAPPIRABLEA-LPGVNLTDVNVLMDCM 231
Db 201 NTLHNDLCTAFEGEPPYITIGDDAQYTLSTFAGPIRTARNANLPGANLTADTVALMDLC 260
QY 232 PFTVARTS-----DATQLSPFCDLFTADEMO-YDYLSQL-KYYGYGAGNPLGP 278
Db 261 PFTVASSSDPATADAGGNGRPLSPFCRLFSSEWRAYDYLSQVKGWYGPNGNPLGP 320
QY 279 AQGVGF-NELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFFAIG 337
Db 321 TQGVGEVNEELLARLGVPRDGTSTNRTLDGDPRTFPLGRPLYADFSDHNDMMGLVGLAG 380
QY 338 LYGTKPLSTTS-VESETIDGYAASVTPFAARAYVEMMOCEAGGGGGBG-----EKEP 390
Db 381 AYDGVPLDKTARRDEPEELGYAASVAVPFAARIYVEKMRCSGGGGGGGGBGROEKDEE 440
QY 391 LVRVLVNDVRVPLHCGGVKDLGRCKLDDFVEGLSFARSNGNNAECFA 437
Db 441 MVRVLVNDVRVMTLKGCGADRGCTLERFTESMAFARGNGKWDLCFA 487

RESULT 9
ID 074677 PRELIMINARY; PRT; 442 AA.
AC 074677;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE (EC 3.1.3.2).
GN PHO1.
OS Pichia angusta (Yeast) (Hansenula polymorpha).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Pichia.
OX NCBI_TaxID=4905;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC34436;
RX MEDLINE=98386672; PubMed=9720203;
RA Phongdara A., Merckelbach A., Keup P., Gellissen G., Hollenberg C.P.;
RT "Cloning and characterization of the gene encoding a repressible acid
phosphatase (PHO1) from the methylotrophic yeast Hansenula
polymorpha."
RT Appl. Microbiol. Biotechnol. 50:77-84(1998).
DR EMBL; AF051161; AAC82537.1; -.
DR HSSP; P34755; IQFX.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid.phosphat. 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 442 AA; 49370 MW; 7087D91A85B053C1 CRC64;

```

```

Db 101 OCQYNGPMVEVNDYEFFVSNKYEDQLTNSDVPSPYAGAKTAQHLGKYIAYNGDLF 160
QY 123 RNIVPFVRASGSDRVIAAEKIFGFSAKLADPAHQASPVINV---IIEG--SGYNNT 177
Db 161 SDSNP-VFTSSGRV-----HOTAKYVVSLEELDQLDQLIQENETSGANSL 209
QY 178 LDHGLCTAFEDSTLGDAAEAFNFTAVFAPPTRAR-LEALPGVNLT--DEDVNVNMDMCPFD 234
Db 210 TPADSCMTY-NGDLGDEYFENATLPYLTLDKNNWKKNSNLNLTLEHDDLELIVDMCAFE 268
QY 235 TVARTSATQSPDCDLFTADEW--QYDYLQSL-KYGYGAGNPL-GPAQGVGFNELIARL 291
Db 269 TNVKGSSAV-----CDLFRNDLVAYSYYANNVYRRGAGNPMSPNPGSVLVNYSYNLL 323
QY 292 THSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTWVSIFFALGLY-NGTKPLSTTSV 350
Db 324 TOADELND-----KVLFSFSDHTDIOQFISALGLIDNGVTEYSILDQV 365
QY 351 --ESIETDGYAASWTVPFAARAYVEMMQCAGGGGEGEKEPLRVLVNDRVVPPLHGGCV 408
Db 366 DFQNIQ----QLSWVTMPMGRIIFTEKLK-----GNASVRYIINDVVIIPVPGCTS 412
QY 409 DKLGRCKLDDF-----VEGLSFARKS 428
Db 413 GPGFSCPIEDFDYITNRLNGIDYVSS 439

RESULT 11
060172 PRELIMINARY; PRT; 463 AA.
AC 060172;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE PROBABLE ACID PHOSPHATASE (EC 3.1.3.2).
GN SPBC21H7.03C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lyne M., Rajandream M.A., Barrell B.G., Xiang Z., Aves S.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; AL023286; CAA1863.1; ..
DR HSSP; P34755; 10FX.
DR InterPro; IPR000560; ..
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS-ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS-ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein.
FT ACT SITE 69
FT ACT SITE 340
FT ACT SITE 340
FT CARBOHYD 98
FT CARBOHYD 98
FT CARBOHYD 104
FT CARBOHYD 104
FT CARBOHYD 221
FT CARBOHYD 221
FT CARBOHYD 324
FT CARBOHYD 324
FT CARBOHYD 439
FT CARBOHYD 439
FT CARBOHYD 458
FT CARBOHYD 458
SQ SEQUENCE 463 AA; 52758 MW; 6C41AF422C6D624A CRC64;

```

Query Match 12.3%; Score 285; DB 3; Length 463;
 Best Local Similarity 24.9%; Pred. No. 4.2e-15;
 Matches 113; Conservative 69; Mismatches 184; Indels 88; Gaps 19;
 QY 23 YSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARYPTS--SKSKKYSA-----L 70

```

Db 43 HEFYFDGLDSA-----FPETCEIQVHLLHGRSNPTGDTVATDVIYSOYLNNEQEKLL 97
QY 71 IERIKNATF-KGKYAFLKTYNTL---GADDLTPFGENQMVNSGIKYYRYRKALARNIV 126
Db 98 NSGIPVNFYSYPENPLCFIKQWTFPVIDAENADQLSSRGLRLEFDLGRQLYQRYKLFDSV 157
QY 127 PFVRASGSDRVIAAEKIFGFSAKLADPAHQASPVINVIIPEG--SGYNTLHDGLCT 184
Db 158 YDINTAEQERVVESAKWTFYGLFGDKMTEKN-----FILISGEKAAGANSLSYNACP 211
QY 185 AFEDSTL-----GDDAEANFTAVFAPPTRARLEAL--PGVNLTDDEVDVMDMCPDPTVAR 238
Db 212 VFEDNFHKNKATDAHAHVNRNIFIEPIVNLAKYFDSSYKLTINDVRSUFYICEYIAIK 271
QY 239 TSDATQLSPDCDLFTADEW--QYDYLQSL-KYGYGAGNPL-GPAQGVGFNELIARLTHSPVQ 297
Db 272 DH-----SDFCSIFTPSEFLNFEYDSDLD-QAYGG-----PVS 304
QY 298 DHTST-----NHTLDS--NPATFPLNATLYADFSDHNTWVSIFFALGLY---NGTKPLS 346
Db 305 EWASTLGGAYINNLADSLRNVTNPDQDKVFLAETHDSNIIPVEAALGFFPDITPQPLP 364
QY 347 TTSVESIETDGYAASWTVPFAARAYVEMMQCAGGGGEGEKEPLRVLVNDRVVPPLHGC 406
Db 365 TD--KNIIYTSOKTSSSFVPFAGNLITELFFC-----SDSKYYVRHLVNOQVYPLDC 414
QY 407 GVDKIGR-----CKLDDFVEGLSFARKSGCNMAECF 436
Db 415 GYGPSGTSGLCELOAYLNSPIRANSTNSGISIF 448

RESULT 12
Q9VV72 PRELIMINARY; PRT; 460 AA.
AC Q9VV72;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE MIPPI PROTEIN.
GN MIPPI OR CG4123.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghavani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin C.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```

```

[1]
RN      SEQUENCE FROM N.A.
RP      Chi H., Wang J., Romano P.R., O'Keefe R.J., Puzas J.E., Rosier R.N.,
RA      Reynolds P.R.;
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RR      ENBL: AF046913; RAD02436.1; -.
DR      FlyBase; FBgn0026061; Mipp1.
DR      InterPro; IPR000560; -.
DR      Pram; PF00328; acid_phosphat; 1.
SQ      SEQUENCE 467 AA; 53602 MW; 1A0571B9958B2632 CRC64;

Query Match      8.4%; Score 196; DB 5; Length 467;
Best Local Similarity 23.5%; Pred. No. 8e-08;
Matches 106; Conservative 59; Mismatches 173; Indels 114; Gaps

Qy      38  DVPK-----GCRVTFVOVLSRHGARYPTSSKKYSALIE-----RIQNA 78
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      46  DIDKOYLVPGCQPKMKWIFHRHGTRELPPKSKMINKASRVAELRDLIINNYQVARTKETDA 105
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      79  TFKKYAFLKTYNTLGCADDLTDPFGNQMVNSCIKPYRYKALARNIVPFV----- 129
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      106  LCQTDLIAIKLWKWN-----SSITPDMEEYLTAAQYEDLRGTAKLYQRYPTVLTANYNDY 162
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      130  ---RASGSDRVIASAEKFIEG-FQSAKLADPAH-----QASPVINW-IIPEG 172
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      163  YQFRTDQRTTESFKAFASLEGFSQNAHPVELPKODLLLRPYDYCSSFKNMYKDEGS 222
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      173  GYNNTLDHGLCTAFEDSTLGDAAENFTAVFAPIRARLEALPGVNLTDDEVVNLMDMP 232
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      223  EYVK-----PHOSKLYNDTIAD-----ISTRGLF--YTLSEADIKLYDMCR 263
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      233  PDTVARTSDATOLSPFCDLTADB-WOYDIQSLKYIYGYGAGNLPAGQVGFNELIARL 291
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      264  YE---QAWNVDNRNSWCACFLPEQITVFEYLEDIKYY-YGS-----GYGFPE-NAHL 310
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      292  THSPVQD---HTSTNHTLDSNPATFPLNATLYADFSDNTMTWSIFFALGLYNGTRPLSTT 348
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      311  NCRLVODLLTHLSN-----PVSPHWVHFHGSTGLTLTALGXKDDIKLRAD 359
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      349  SVESITETDGYAASWTVPFAARYVMQMCCAGGGGGGKEPELVRVLVNDVRVPLHGCVG 408
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      360  NYDSITSLRRKWSLLIDPPAANFVAKVLDLPA-----DLDRKVV-FFLNQQAOLDWCSV 413
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      409  DKLGRCKLDDFEV-----GLSFARSGG 430
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      414  ---GLCKNSDVLEKYKTIADACGEYVCRGG 442
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
O96420  PRELIMINARY;      PRT;      453 AA.
AC      O96420;
DT      01-MAY-1999 (TRENBLrel. 10, Created)
DT      01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT      01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE      MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 2.
GN      MIP2 OR CG4317.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_Taxid=7227;
RN      SEQUENCE FROM N.A.
RP      Chi H., Wang J., Romano P.R., O'Keefe R.J., Puzas J.E., Rosier R.N.,
RA      Reynolds P.R.;
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RR      ENBL: AF046912; RAD02435.1; -.
DR      FlyBase; FBgn0026060; Mipp2.
DR      InterPro; IPR000179; -.
DR      PROSITE; PS00193; CYTOCHROME_B_QC: UNKNOWN.1.
SQ      SEQUENCE 453 AA; 52799 MW; 9D03CCDCEBB8B6B CRC64;

```

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nayak K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo R., Pittman G.S., Pan S.K., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupsis M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
DR ENBL; AE003527; AAF49450.1; -;
DR FlyBase; FBgn0026061; Mippl.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid_phosphat; 1.
SQ SEQUENCE 460 AA; 52685 MW; F00B25718E40807D CRC64;

Query Match 8.9%; Score 207.5; DB 5; Length 460;
Best Local Similarity 24.2%; Pred. No. 9e-09;
Matches 109; Conservative

QY 38 DVPK-----GCRVTFVQVLRSRHGARYPSTSSKSKYSALIE-----RIQKNA 78
DB 46 DIDKOYLPGQPQOKMWIFHRHGRTRLPKKSINKASRVAELRDLIINNYQVARTKPETDA 105
QY 79 TFKGKIAYFLKYNTYLGLADDTLTPGENOMVNSG-----IKFYRYKALARNIVPFVRA 131
DB 106 LCQTDLIAIKLKWN---SSTIPDMEEYLTAQGVEDLRGTAKLYORYPTV-----LTA 156
QY 132 SGSD----RVIASAEKFTEG-FQSACLADPAH-----QASPVINV-IIPEGSG 173
DB 157 NYNDIYYORTTESFAFAGLEFGSONAAHPVEIPKODLLRPDYCSFKKNVNYKDEGSE 216
QY 174 YNNTLDHGICTAFEDSTLGDDAENFTAVAPPTRARLEALPGVNLTDEDVVNLMDMPF 233
DB 217 YYK-----FHQSKLYNLTAD-----ISTRLGFL--YTLEEADIKLMYDMCKY 257
QY 234 DTIVARTSATQLSPCDLFTADE-WQYDYLOSLKYGYGAGNPGLPACQGVGFNELIARLT 292
DB 258 E---QAWNVDNRNVWCAGFLPEQITVFYELEDKYY-YGS-----GYGPPE-NAHLN 304
QY 293 HSPGVQ---HTSTNTHLSNPATFPNLATLYADFSDHNTMWSIFFALGIYNCTKPLSTS 349
DB 305 CRLVQDLLTHLSN-----PVSPHVAAHFHSHTGTLTLLTALGIQKDIDKLRADN 353
QY 350 VSIETFDGVAASWTVPFAARATVEMMOCEAGGGGEKEPELVRVLNDRVVPLHCGCYVD 409
DB 354 YSLTSRRNKSSLIDPPFAANFVAVKYDCFA-----DLDRKKV-VFLNQQAQVOLDWCVS- 406
QY 410 KLGRCKLDLDFE-----GLSFARSGG 430
DB 407 --GLKWSDVLEKYKTADACGEYCRTGG 435

RESULT 13
O96421 PRELIMINARY; PRT: 467 AA.

ID AC O96421
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-NAR-2001 (TREMBLrel. 16, Last annotation update)
DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 1.
GN MIPPI OR CG4123.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.

Query Match		8.3%;	Score 192.5;	DB 5;	Length 453;
Best Local Similarity		23.4%;	Pred. No. 1.5e-07;		
Matches	101;	Conservative	64;	Mismatches	174; Indels 93; Gaps

OY 24 SPFFSLA--DESAIPDPKGCGRVTFVOVLRHGARYPTSSKKYSALIERIQKNATFK 81
 :.:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 43 TYRAIANYE---TPPYACCHPRITWTTIRHGTNRPSVESVIIQAQNRILSEIKRILDQ 99

OY 82 KYAF-----LKTNYVT-LGADD----LTPFGENQMVSNGIFRYRYKALARNVIP-- 127
 |.:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 100 TKPICTABELEKLROWHMHLNATEDEKLLVAEGEDELIELAEERMQRFPDLPPELYNPE 159

OY 128 --FVRASGSDRVIASEAEIFEG-----FSQAKIADPAHQASPVINVIPEGSGYNNTLD 179
 :.:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 160 WYIFYKTATORTLKAAESFAFGLEGRRIHVTRYPPLHE-DPVLREY--KGCGKWKT-- 214

OY 180 HGLCTAFEDSTLGDDAEANFTAVFA-PPIRARLE-----ALPGVNLTDEDVYNLMDCP 232
 :.:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 215 -----DVDKNPBTLLNNARRELAEPMQSASVGQVRSSFLRP--DLQPEDVOLMTTVCA 264

OY 233 EDTVART-----SDATOLSPFDLFATAEWQ-YDYLOSLKY--GYGAGNLPGPAQG 282
 :.:|.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 265 FETAHWRPHDSGSKSYESVMCFDFDAAMEALEFFEDELLEYMNDGYG----- 313

OY 283 GENELIARLTHSPQQDH-----TSTNHTLDSNPATFPLNATLYADESHDNMTVSIFFALGI 338
 |||.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 314 ---VELTHRIACPATIDMEFAAISSESSEETRQRR-----ANATLY-FHTSGTLLKLAHLGL 364

OY 339 YNGTKPLSTSIESIEDTYGAASWTVPFAPARAYEMMQCEAGGGGEKEPLEIVRLVND 398
 |||.:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.
Db 365 ARDNKPLTHKHFAERL--WRTSOIDAFATINLAFRLRVCDCKG-----NPQVLVHQE 414

OY 399 RVVPLHGCGVDK 410
 |||||.:.|.|.|.|.|
Db 415 RVVRLPGCPQDK 426

RESULT	15
ID	Q9W438
NCBI TaxId	PRELIMINARY; PRT; 453 AA.
ACC	O9W438;
DT	01-MAY-2000 (TEMBLrel. 13, Created)
DT	01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT	01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE	MIPP2 PROTEIN.
OS	MIPP2 OR CG4317.
GN	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX	Ephydroidea; Drosophilidae; Drosophiliinae; Drosophila.
NCBI_TaxId=7227;	[1]
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
MDLINE=20196006;	PubMed=107311132;
RA	Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Van K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bereman B.P., Bhattacharya D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brocksteins P., Brottier P.,
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doull L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.E., Ferraz C., Ferreira S., Fleischmann W.,
RA	Foster C., Gabrielian A.C., Garg N.S., Gelbart W.M., Glasser K.,

Search completed: October 26, 2001, 15:17:25
Job time: 3976 sec

us-09-488-265-27.rspt

Sat Oct 27 15:25:44 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:03 ; Search time 57.24 Seconds
(without alignments)
462.835 Million cell updates/sec

Title: US-09-488-265-27

Perfect score: 2321

Sequence: 1 NSHSCDVTGQYQCPPEISHLW.....DFVEGLSFARSGGNWAECA 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_0601.*
- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2321	100.0	437	21	AA20525
2	2321	100.0	437	21	AA20525
3	2115	91.1	441	21	AA20523
4	2115	91.1	467	21	AA20524
5	2115	91.1	467	21	AA20526
6	2099	90.4	467	21	AA20533
7	2093	90.2	467	20	AA20527
8	2093	90.2	467	21	AA20527
9	2093	90.2	467	21	AA20534
10	2093	90.2	467	21	AA20534
11	2067	89.1	467	21	AA20531

12	2064.5	88.9	424	21	AA20536
13	2064.5	88.9	424	21	AA20536
14	2061	88.8	467	21	AA20532
15	2044	88.1	467	20	AA20526
16	2044	88.1	467	21	AA20526
17	2044	88.1	467	21	AA20526
18	2020	87.0	441	21	AA20514
19	2020	87.0	467	20	AA20514
20	2020	87.0	467	20	AA20515
21	2020	87.0	467	21	AA20515
22	2020	87.0	467	21	AA20515
23	2014	86.8	467	20	AA20515
24	2013	86.7	467	20	AA20515
25	2013	86.7	467	20	AA20515
26	2005	86.4	467	20	AA20515
27	2004	86.3	467	20	AA20515
28	1997	86.0	431	21	AA20515
29	1997	86.0	431	21	AA20515
30	1954	84.2	467	21	AA20530
31	1954	84.2	467	21	AA20530
32	1922	82.8	467	20	AA20530
33	1922	82.8	467	21	AA20530
34	1919	82.7	467	21	AA20530
35	1797.5	77.4	467	21	AA20529
36	1777.5	76.6	467	20	AA20529
37	1777.5	76.6	467	21	AA20529
38	1777.5	76.6	467	21	AA20529
39	1762.5	75.9	440	21	AA20529
40	1762.5	75.9	465	19	AA20529
41	1762.5	75.9	465	20	AA20529
42	1759.5	75.8	440	21	AA20529
43	1759.5	75.8	440	21	AA20529
44	1754.5	75.6	440	21	AA20529
45	1754.5	75.6	440	21	AA20529

ALIGNMENTS

RESULT 1
ID AAB20525 standard; Protein; 437 AA.
XX
AC AAB20525;
XX
DT 05-DEC-2000 (first entry)
XX
DE Consensus phytase 11 SEQ ID NO:27.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure.
XX
OS Synthetic.
XX
PN WO200043503-1
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-DK00025.
XX
PR 22-JAN-1999; 99DK-0000092.
XX
PR 21-SEP-1999; 99DK-0001340.
XX
PA (NOVO) NOVO NORDISK AS.
XX
PI Lehmann M;
XX
DR WPI; 2000-491161/43.
XX
PT Novel phytases with improved properties such as temperature stability,

Consensus phytase
Initial consensus
Consensus phytase
Consensus phytase
Consensus phytase
Mutant phytase-1,
Consensus phytase
Ascomycete consens
Fungal phytase pro
Consensus phytase
Phytase-1, a conse
Fungal consensus p
Fungal consensus p
Fungal consensus p
Fungal consensus p
Consensus phytase
Initial consensus
Consensus phytase
Conspyl12, a deriv
Consensus phytase-
Phytase-7, a deriv
Consensus phytase
Mutant Aspergillus
A. fumigatus phyta
Aspergillus fumiga
Mutant Aspergillus
Aspergillus fumiga
Aspergillus fumiga
A. fumigatus phyta
Aspergillus fumiga
Aspergillus fumiga
Aspergillus fumiga
Aspergillus fumiga

RESULT	2	
AAAY69567		standard; protein; 437 AA.
XX		
XX		
AC		
XX		AAAY69567;
XX		
DT		19-APR-2000 (first entry)
XX		
XX		Phytase-11, a consensus mature phytase.
DE		
XX		
XX		Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
KW		thermostable; animal feed; monogastric animal; phytate phosphorus;
KW		phosphate availability; consensus; phytase-11.
KW		
XX		

```
SQ      Sequence   437 AA;
Query Match          100.0%; Score 2321; DB 21; Length 437;
Best local Similarity 100.0%; Pred. No. 2.e-218;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 N H S C D T V D G I Q C P E I S H L W C Q Y S F P F S L A D E S A I S P D V P K G R V T F V O L S R H G A R Y P T    60
```

Db 1 nshscdtvdygcpcishlwgqyspfsladesaispdkgrvtfvqlsrhgarypt 60
 QY 61 SSKSKYSALIERIKNATFKGKXAFKTYNTLIGADDLTPFGENQMVNSGKIFVRRYKA 120
 Db 61 sskskysalieriknatkfkgkafktyntylgaddltfpgenqmvnsqkifvrryka 120
 QY 121 LARNIVPFVRASGSDRVIASAEKFTGFSQAKLADPAHQASPVINVIIEGSGYNNTLDH 180
 Db 121 larnivpfvrasgsdrviasaeekflegfqsakladpahqaspviniiepgsgynntldh 180
 QY 181 GLCTAFEDSTGLDDEANFTAVFAPPFARLEALPGVNLTDDEVVNLMDCPFDVARTS 240
 Db 181 glctafedstlgddaeantafavappirarlealpgvnltdedvvnlmcpfdvarts 240
 QY 241 DATQLSPFCDLFTADEWYDYLQSLKYGYGAGNPLGPAQGVGFNELLARLTHSPVQDHT 300
 Db 241 datqlspfcldftadewydyqlskkygygagmplgagvgvgnellarlthspvqdh 300
 QY 301 STNHTLDSNPATFPLNATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSVESIETGYAA 360
 Db 301 stnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsvesietgyaa 360
 QY 361 SWTPFAARAYVEMMOCEAGGGGEGEKEPLVRVLVNDRVVPLHGCVDKLGCKLDDFV 420
 Db 361 swtpfaarayvemmoceagggggegkeplvrvlvndrvvplhgcgvdklgckldfv 420
 QY 421 EGLSFARSGGNWAECEFA 437
 Db 421 eglfsarsggnwaecefa 437

RESULT 3
 AAB20523
 ID AAB20523 standard; Protein: 441 AA.
 AC AAB20523;
 DT 05-DEC-2000 (first entry)
 XX Consensus phytase 10 (Fcp10) SEQ ID NO:24.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.
 OS WO200043503-A1.
 XX 27-JUL-2000.
 XX 21-JAN-2000; 2000WO-DK00025.
 XX 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PA Lehmann M;
 PI WPI; 2000-491161/43.
 DR Novel phytases with improved properties such as temperature stability,
 XX pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX Example 2; Fig 4a-d; 240pp; English.
 PS The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.

CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX Sequence 441 AA;
 SQ Query Match 91.1%; Score 2115; DB 21; Length 441;
 Best Local Similarity 93.5%; Pred No. 3.1e-198;
 Matches 419; Conservative 2; Mismatches 9; Indels 18; Gaps 10;
 QY 1 NSHSCDTVD-GYQC-PEISHLMGQYSPFSLADESAISPDKGRVTFVQVLSRHGARY 58
 Db 1 nshscdtvdygcpcishlwgqyspfsladesaispdkgrvtfvqlsrhgary 60
 QY 59 PTSSSKYSALIERIKNAT-FKGKXAFKTYNTLIGADDLTPFGENQMVNSGKIFVRR 117
 Db 61 ptssskysalieriknatafkgyafktyntylgaddltfpgenqmvnsqkifvrr 120
 QY 118 YKALARNIVPFVRASGSDRVIASAEKFTGFSQAKLADPA---HQASPVINVIIEGSGY 174
 Db 121 ykalarlivpfvrasgsdrviasaeekflegfqsakladpahqaspviniiepgsgy 180
 QY 175 NNTLDHGLCTAFEDSTGLDDEANFTAVFAPPFARLEA-LPGVNLTDDEVVNLMDCP 233
 Db 181 nntldhglctafeseldgdeantafavappirarleahpgvnltdedvvnlmcpf 240
 QY 234 DTVARTSDATQLSPFCDLFTADEW-QYDYLOSL-KYCYGAGNPLGPAQGVGF-NELLAR 290
 Db 241 dtvartsdatqlspfcldftadewiqyqlskkygygagmplgagvgvgnellar 300
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFALGLYNGTKPLSTTSV 350
 Db 301 lthspvqdhstnhtldsnpatfplnatlyadfsdntmvsiffalglyngtkplsttsv 360
 QY 351 ESI-ETDCGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRVLVNDRVVPLHGCVD 409
 Db 361 esieetdgyaaswtvpfaarayvemmoce-----ekeplvrvlvndrvvplhgcgv 413
 QY 410 KLGCKLDDFVEGLSFARSGGNWAECEFA 437
 Db 414 klgrckrddfveglfsarsggnwaecefa 441

RESULT 4
 AAB20524
 ID AAB20524 standard; Protein: 467 AA.
 AC AAB20524;
 DT 05-DEC-2000 (first entry)
 XX Consensus phytase 10 SEQ ID NO:26.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.
 OS WO200043503-A1.
 XX 27-JUL-2000.

to the feed. The level of phosphate pollution in the environment is reduced by adding phytase to animal feed, as the animal can make use of the inorganic phosphate liberated from phytate phosphorus using the enzyme. The phytase formulation of the invention has an improved thermostability and can therefore remain stable during long-term storage and can withstand feed processing methods such as extrusion, expansion and pelleting. The present sequence represents a consensus sequence, designated phytase-10, which was derived from the mature phytase sequences from a variety of fungi (AAV69544-Y69546, AAV69548-Y69556, AAV69564) and the Basidiomycetes phytase consensus AAV69563 and additionally contains the *Aspergillus terreus* cbsl16.46 signal peptide at the N-terminus.

Sequence 467 AA;

Query Match 91.1%; Score 2115; DB 21; Length 467;
Best Local Similarity 93.5%; Pred. No. 3.4e-198;
Matches 419; Conservative 2; Mismatches 9; Indels 18; Gaps 10;

QY	1	NSHSCD	TVD - GYOC - PEISHLWGOYSPFFSLADESAISPDVPKGRVTFQVLSRHGARY	58
Db	27	nshscd	tdvgdygcfpeshlwgyospfsladesaispvpkgcrvtfvqlstrngary	86
QY	59	PTSSKSKYSALIERIQKNAT - PKGKTAFTKTYNTLGDLLTPFGENOMVANSIGIKFYRR	117	
Db	87	ptsskskysalialkaqkna tafkgfyafktyntylgaddltpfgeqgmvnsgikfyrr	146	
QY	118	YKALARNIVPVRASGSDRVIASAEKTECFQSAKLADPA - -- HQASPVINVIIEGSGY	174	
Db	147	ykalkarvipvrasgsdrviasaekfiegqakladpganphgaspvinvilipagay	206	
QY	175	NNTLDHGLCTAFEDSTLGDDAENFTAVAPPTRARLEA - LFCVNIITDEDVNLMDMCPF	233	
Db	207	mntldhglctafesegldgdeanfnavfappirarleahlpvnltdedvnlmdmcpf	266	
QY	234	DTVARTSDATQSPCDFLETADEW - OYDYLOSH - KYGYGACNPLGPAGGVGF - NEILIAR	290	
Db	267	dtvartsdatqspcdfldhewiaydyqlsgkygygagnpilgpaqgvgfneiliar	326	
QY	291	LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHDNTMVSIFFAALYNGTKPLSTTSV	350	
Db	327	lthspvqdhtstnhtldsnpatplnatlyadfshdntmvsiffaalyngtkplsttsv	386	
QY	351	ESI - ETDCGYAASWTVPFAARYVMQCEAGGGEKEPLRVLVNDRVVPVPHGCCVD	409	
Db	387	esieetdgyaaswtvpfaaryvmmqcea ----- ekeplrvlvndrvvpvhgcvgd	439	
QY	410	KLGRCKLDDRFVGLSFARSGGNWAECEFA	437	
Db	440	klgrckrddrfvglsfarsggnweecfa	467	

RESULT 6

AAB20533
 ID AAB20533 standard; Protein; 467 AA.

AC AAB20533:

DT 05-DEC-2000 (first entry)

Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95

Phytase; mutant; thermostability; mutation; mutagenesis; pH stability; temperature stability; pH profile; temperature profile; reaction rate; specific activity; substrate specificity; substrate cleavage pattern; substrate binding; position specificity; phytate degradation rate; food; feed; phytate; manure

OS Synthetic.

WO200043503-A1.

XX

~~PD-27-JUL-2000-~~

21-JAN-2000; 2000WO-DK00025.

XX 2
XX 2

PR 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.
XX

PA (NOVO) NOVO NOROISY 100

FA (NOVO) NOVO NORDISK AS.
XX

PI Lehmann M.

XX
LEIGH M.,

DR WPI: 2000-491161/43

DR N-PSDB; AAA73292.

[illegible]

PT Novel phytases with improved properties such as temperature stability,
PT pH stability and substrate specificity, for use in pharmaceuticals and
PT compound foods and feeds -

PS Disclosure; Fig 24a-c; 240pp; English.

The present invention describes improved phytases, preferably with increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. thermostability.

temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence represents a phytase sequence from the present invention.

AA	Sequence	467	AA:
SQ			

Query Match	90.4%;	Score 2099;	DB 21;	Length 467;
Best Local Similarity	92.6%;	Pred. No. 1.2e-196;		

Matches	415;	Conservative	4;	Mismatches	11;	Indels	18;	Gaps	10;
---------	------	--------------	----	------------	-----	--------	-----	------	-----

QY 1 NSHSCDSTD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPKGCRTVTFVQVLSRHGARY 58

Db 27 nshscdtvdgggyqc:peishlwtyspffsladesaisdpvkgrvtfvqvlshrghary 86

QY 59 PTSSKSKYSALIERIQKNAT-FKGKYAFKTYNTILGADDITPFGENQMVNSGKIFYRR 117

Db 87 ptsskksaysalieaiqnatafkgkyafltktyntylgadtaltpfgeqgmvsngikfyrr 146

QY 118 YKALARNIVFVRASGSDRYASAEKFIEGFQSAKLADPA---HQASPVINVIIEGSGY 174

Db 147 ykalarxivfirasgsdrviasaekfieqfakladpaganhqaspviniipagay 206

QY 175 NNTLDHGLCTAFEDSTLGDDAEANTAVFAPTRARLEA-LPGVNLTDDEVVNLMDMCPF 233

Db 207 nntldhglctafeestlgddveanftavfappirleahlpvgvnltdedvnlmdmcpf 266

Qy 234 DTVARTSDATQLSPPFCDIIFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290

Db	267	dtvartsdatsqspfdlftdhewiqdyqlslgkygygagnlqpaqgygvnlliar	326
----	-----	---	-----

Qy 291 LTHSPQDHTSTNHTLDSNPATFPPLNATLYADESHDNTMVSIFFALGLYNGTKPLSTTSV 350

Db 327 lthspvqdhirstnhtldsnpatfpinatlyadfsdhntmvsiffalglyngtkplsttsv 386

QY 351 ESTDGYAASWTVPFAARAYVEMOCEAGGGGEGEKEPLRVRLVNDRVVPLHGCGVD 409

Db 387 esieetdgysaswtvpfaarayvenmqcea-----ekeplrvrlvndrvvplhgcavd 439

QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437

Db
440 klgrckrddffveglsfarsgggnweecfa 467

Db 267 dtvartsatqlspfcldlfthdewiqdyqlslokygygagpnlpqagvgvfnelliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTWTSIFFALGILYNGTKPLSTTSV 350
 Db 327 lthspvqdhstnthtldsnpatfnatlyadfshtntmvsiffalglyngtkplstsv 386
 QY 351 ESI-ETDGYAASWTVPFAARAYVENMOCEAGGGGEGEKEPLVRLVNDVRVPLHGGVD 409
 Db 387 esieetdgysaswtvpfaaryvemumqcea-----ekeplvrvlndrvrvplhggvd 439
 QY 410 KLGCRKLDLDFVEGLSFARSGGNWAECA 437
 Db 440 klgcrkrddfvglsgfarsggnweecfa 467

RESULT 8

AAB20527
 ID AAB20527 standard; Protein: 467 AA.

AC AAB20527;

XX 05-DEC-2000 (first entry)

DT Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.

DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 XX temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

XX W0200043503-A1.

PN 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI: 2000-491161/43.

DR N-PSDB; AAB73234.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 4; Fig 8a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, the velocity and level of
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

SQ

RESULT 7

AAAY43170
 ID AAY43170 standard; Protein: 467 AA.

XX AC AAY43170;

XX 06-JAN-2000 (first entry)

DT Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.

DE Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 KW consensus sequence.
 KW Synthetic.

OS W09948380-A1.

PN 30-SEP-1999.

XX 22-MAR-1999; 99WO-DK00154.

XX 23-MAR-1998; 98DK-0000407.

PR 19-JUN-1998; 98DK-0000806.

PR 18-SEP-1998; 98DK-0001176.

PR 22-JAN-1999; 99DK-0000091.

XX 22-JAN-1999; 99DK-0000093.

XX (NOVO) NOVO-NORDISK AS.

PA Petersen S;

XX WPI: 1999-591030/50.

DR N-PSDB; AAZ31521.

XX Preparing animal feed using a thermostable phytase

XX Example 3; Fig 10; 71pp; English.

XX This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently. In addition to improved
 CC phytase-expressing transgenic plants, these plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

XX Query Match 90.2%; Score 2093; DB 20; Length 467;

XX Best Local Similarity 92.4%; Pred. No. 4.8e-196;

XX Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;

QY 1 NSHSCDFVD-GYOC-PEISHLWGOYSPFFSLADESAISPDVPKGRVTFVQVLSRHGARY 58

Db 27 nshscdctvdggyqcfpeishlwgtyspffsladesaispdvpgkrvtfvqlsrhgary 86

QY 59 PTSSSKKYSALIERIOKNAT-FPGKYAFUKTYNYTLGADDLTPFGENQMVNSGKIFYRR 117

Db 87 ptssaskaysallcaiqknatafkgyafiktyntylgaddltpfgeqmvnsgkifyrr 146

QY 118 YKALARNIVFVRASGDRVIASKEKIEGFSQAKIADPA---HQASPVINVIIPGSGY 174

Db 147 ykalarlivfirasgdrviassekiefqskladpaganphqaspvinviipgagay 206

QY 175 NNTLDHGLCTAFEDSTLGDAAEAFTAVFAPPRIARLEA-LPGVNLTPDEDVNLMDMCPF 233

Db 207 nntldhglctafecselgddeanftavfappirarleahlpvgnltddedvnlmdmcpf 266

QY 234 DTVARTSDATOLSPFCDLFTADEW-QYDYLOSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290

Query Match 90.2%; Score 2093; DB 21; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLMGQYSPFFSLADESAISPDPVKGCRVTFVOVLSRHGARY 58
 DB 27 nshscdvtvggvcqfpeishlwgtyspffsladesaispdpvkgcrvtfvqvlshrgary 86

QY 59 PTSSSKKYSALIERIOKNAT-FKGYAFKTYNTLGAADLTTPGEMQVNSGKIFYRR 117
 DB 87 ptssaskaysalieaiaqknatafkgyafktyntlgadlttppgeqmqvnsqikfyrr 146

QY 118 YKALARNIVPVRASGSDRVIASAEKFTIEGFSQAKLADPA---HQASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaeekfiegfsqakladpaganphqaspinviipegagy 206

QY 175 NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPPIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafeesigddveanftavfappirarleahpgvnltdedvnlmdmcpf 266

QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatsqlspfcldlthdewiqdyqlsgkygygagnpplgpaqgvgfveliar 326

QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFPFGALYNGTKPLSTTSV 350
 DB 327 lthspvqgdhtstnhtlgsnpatfpnlntlyadfsdntmvsiffalglyngtkplsttsv 386

QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEGEKEPLVRLVNDVRVPLHGGCVD 409
 DB 387 esieetdgyaswtvpfaarayvemmqcea-----ekeplvrvlvndrvrplhgcgvd 439

QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 DB 440 klgrckrdddiveglstarsgnwceefa 467

RESULT 9

AAB20534
 ID AAB20534 standard; Protein; 467 AA.
 XX
 AC AAB20534;
 XX
 DT 05-DEC-2000 (first entry)
 DE
 DE Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.
 XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX
 OS Synthetic.
 XX
 XX WO200043503-A1.
 XX
 XX 27-JUL-2000.
 XX
 XX 21-JAN-2000; 2000WO-DK00025.
 XX
 XX 22-JAN-1999; 99DK-0000092.
 XX
 XX 21-SEP-1999; 99DK-0001340.
 XX
 XX (NOVO) NOVO NORDISK AS.
 XX
 XX Lehmann M;
 XX
 XX WPI; 2000-491161/43.
 XX
 XX N-PSDB; AAA73293.
 XX
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and

compound foods and feeds -
 Disclosure; Fig 25a-c; 240pp; English.

The present invention describes improved phytases, preferably with increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and end level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence represents a phytase sequence from the present invention.

Sequence 467 AA;

Query Match 90.2%; Score 2093; DB 21; Length 467;
 Best Local Similarity 92.4%; Pred. No. 4.8e-196;
 Matches 414; Conservative 4; Mismatches 12; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLMGQYSPFFSLADESAISPDPVKGCRVTFVOVLSRHGARY 58
 DB 27 nshscdvtvggvcqfpeishlwgtyspffsladesaispdpvkgcrvtfvqvlshrgary 86

QY 59 PTSSSKKYSALIERIOKNAT-FKGYAFKTYNTLGAADLTTPGEMQVNSGKIFYRR 117
 DB 87 ptssaskaysalieaiaqknatafkgyafktyntlgadlttppgeqmqvnsqikfyrr 146

QY 118 YKALARNIVPVRASGSDRVIASAEKFTIEGFSQAKLADPA---HQASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaeekfiegfsqakladpaganphqaspinviipegagy 206

QY 175 NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPPIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafeestlgddveanftavfappirarleahpgvnltdedvnlmdmcpf 266

QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatsqlspfcldlthdewiqdyqlsgkygygagnpplgpaqgvgfveliar 326

QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFPFGALYNGTKPLSTTSV 350
 DB 327 lthspvqgdhtstnhtlgsnpatfpnlntlyadfsdntmvsiffalglyngtkplsttsv 386

QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGEGEKEPLVRLVNDVRVPLHGGCVD 409
 DB 387 esieetdgyaswtvpfaarayvemmqcea-----ekeplvrvlvndrvrplhgcgvd 439

QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
 DB 440 klgrckrdddiveglstarsgnwceefa 467

RESULT 10

AAY69569
 ID AAY69569 standard; Protein; 467 AA.
 XX
 AC AAY69569;
 XX
 DT 19-APR-2000 (first entry)
 DE
 DE Mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.
 XX
 KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; muten.
 XX
 OS Aspergillus terreus 9A1.

[illegible]

RESULT	11	
AAAB20531		
ID	AAAB20531	standard; Protein: 467 AA.
XX		
XX	AAAB20531;	
XX		
XX	AC	
XX		
XX	05-DEC-2000	(first entry)
DT		
XX		
XX	Consensus	phytase 3 thermo 11 Q50T protein SEQ ID NO:91.
DE		
XX		
XX	Phytase;	mutant; thermostability; mutation; mutagenesis; pH stability;
KW	temperature	stability; pH profile; temperature profile; reaction rate;
KW	specific	activity; substrate specificity; substrate cleavage pattern;
KW	substrate	binding; position specificity; phytate degradation rate;
KW	food;	feed; phytate; manure.
XX		
XX	Synthetic.	
XX		
XX	WO200043503-A1.	
DN		
XX		
XX	27-JUL-2000.	
PD		
XX		
XX	21-JAN-2000;	2000WO-DK00025.
PF		
XX		
XX	22-JAN-1999;	99DK-0000092.
PR		
XX	21-SEP-1999;	99DK-0001340.
XX		
XX	(NOVO)	NOVO NORDISK AS.
PA		
XX		
XX	Lehmann M;	
PI		
XX		
XX	WPI: 2000-491161/43.	
DR		
DR	N-FSD8: AAA73290.	
XX		
XX	Novel	phytases with improved properties such as temperature stability,
XX	pH	stability and substrate specificity, for use in pharmaceuticals and
PT		
PT		

PT compound foods and feeds -

XX Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g. file,
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 89.1%; Score 2067; DB 21; Length 467;
 Best Local Similarity 91.3%; Pred. No. 1.7e-193;
 Matches 409; Conservative 7; Mismatches 14; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PRISHLWQYSPFFSLADESAISPDPKGRVTFVQVLSRHGARY 58

DB 27 nshscdvtvdygqpcpeishlwgtyspyfsladesaispdpkdcrtvfvqlsrhgary 86

QY 59 PTKSKKYSALIERIOKNAT-FKGKYAFILKTYNTLGADLTTPFGENQWNSGKIFRYR 117

DB 87 ptkskkysalieraiknatafkgyafiktyntlgadltfpfgendqmwvnsqkfyrr 146

QY 118 YKALARNIVFVRASGSDRVIAAEKFIQFQSAKLADPA--HOASPVINVIIPGSGY 174

DB 147 ykalar-kivpfirasgsdrviasaekfiqfqsakladpqsqhqaspviniipgsgy 206

QY 175 NNTLDHGLCTAFEDSTLGDDAEANTAVFAPPTRARLEA-LPGVNIJTDDEVVNLMDMCP 233

DB 207 nntldhglctafedstlgsdveanftalpapalarleadlpgvntldedvylmdmcpf 266

QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLSL-KYGYGAGNPLGPAQGVGF-NELTAR 290

DB 267 dtvartsdatspfcalthdewdyqlsgkygygagnp1gpaqgvgfanelar 326

QY 291 LTHSPVQDHTSNHTLDSNPATFPLNATLYADFSDNTMVSIFFALGLYNGTKPLSTTSV 350

DB 327 lthspvqdnstnhtldsnpatfplnatlyadfsdntmismisiffalglyngtkplsttsv 386

QY 351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRVLVNDRVVPLHGCGVD 409

DB 387 esleetdgyaswtvpfaarayvemmqcqa-----ekeplvrvlvndrvvplhgcvd 439

QY 410 KIGRCKLDFVEGLSFARSGGNWAECA 437

DB 440 kigrckrddfveglsfarsggnwaecca 467

RESULT 12

AAB20536

ID AAB20536 standard; Protein: 424 AA.

XX AAB20536;

AC AAB20536;

XX 05-DEC-2000 (first entry)

DE Consensus phytase protein from Fig 4.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO2000043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 93DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 4a-d; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 424 AA;

Query Match 88.9%; Score 2064.5; DB 21; Length 424;
 Best Local Similarity 94.3%; Pred. No. 2.5e-193;

Matches 414; Conservative 0; Mismatches 8; Indels 17; Gaps 12;

QY 1 NSHSCDVID-GYQCPEISHLWQYSPFFSLADESAISPDPKGRVTFVQVLSRHGARY 59

DB 1 nshscdvtvdygqpcpeishlwgtyspyfsladesaispdpkdcrtvfvqlsrhgary 59

QY 60 TSSKSKYSALIERIOKNATFKGYAFILKTYNTLGADLTTPFGENQWNSGKIFRYRYK 119

DB 60 tsskskysali-aiknatfkgyafiktyntlgadltfpfgendqmwvnsqkfyrryk 118

QY 120 ALARNIVFVRASGSDRVIAAEKFIQFQSAKLADPAHOASPVINVIIPGSGYNTLTD 179

DB 119 alar-ivpfvrasgsdrviasaekfiqfqsakladpa-qaspviniipg-gyntld 175

QY 180 HGLCTAFEDSTLGDDAEANTAVFAPPTRARLEALPGVNIJTDDEVVNLMDMCPDFTVART 239

DB 176 hgactafeseldgdveanftavfapptrarlealpgvntldedvnlmdmcpdftva-t 234

QY 240 SDATQSPFCDLFTADEWQYDYLSLKYGYGAGNPLGPAQGVGFNELLARLTHSPVQDH 299

DB 235 sdatqspfcldft-hewqdyqlsgkygygagnp1gpaqgvgfnelar1thspvqdh 293

QY 300 TSTNHTLDSNPATFPLNATLYADFSDNTMVSIFFALGLYNGTKPLSTTSVE-SIETDGY 358

DB 294 tstnhtldsnpatfplnatlyadfsdntmvsiffalglyngt-plsttsvepseetdgy 352

QY 359 AASWTVPFAARAYVEMMQCEAGGGGEGEKEPLVRVLVNDRVVPLHGCGVDKLGCKLDD 418

DB 353 aaswtvpfaarayvemmqce-----ekeplvrvlvndrvvplhgcvdklgck-dd 405

QY 419 FVEGLSFARSGGNWAECEFA 437
 DB 406 fveglstfarsggnweecfa 424
 RESULT 13
 ID AAY69565 standard; protein; 424 AA.
 XX
 AC AAY69565;
 XX
 DT 19-APR-2000 (first entry)
 XX
 DE Initial consensus mature phytase #2.
 XX
 KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus.
 XX

OS Aspergillus terreus 9A1.
 OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger str. NRRL3135.
 OS Aspergillus fumigatus AFCC13073.
 OS Aspergillus fumigatus ATCC32722.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Emericella nidulans
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Paxillus involutus N005693.
 OS Trametes pubescens N09343.
 OS Agrocybe pediades N009289.
 OS Penicillium lyali N006113.
 OS Thermomyces lanuginosa.
 XX

EP969089-A4-
 05-JAN-2000.

23-JUN-1999; 99EP-0111949.
 29-JUN-1998; 98EP-0111960.

(HOFF) HOFFMANN LA ROCHE & CO AG F.
 Brugger R, Lehmann M, Wyss M;
 WPI; 2000-099429/09.

New stabilized enzyme formulation, useful for feed compositions for monogastric animals -

Example 4; Fig 16; 101pp; English.

The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate phosphohydrolase) and one or more stabilising agents including xylitol or ribitol; polyethylene glycols with a molecular weight of 600 to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic, glutaric and succinic acid; carboxymethylcellulose; and sodium alginate. The stabilised phytase formulation is used in a method for preparing a feed composition for monogastric animals (e.g., pigs, poultry) and provides a monogastric animal with its dietary requirements of phosphorus. Although a large amount of phosphate is present in animal feed in the form of phytate phosphorus, monogastric animals are unable to utilise this form of phosphate, resulting in the addition of extra phosphate to the feed of such animals. Phytase enhances the nutritional value of plant material without the need for adding additional phosphate to the feed. The level of phosphate pollution in the environment is reduced by adding phytase to animal feed, as the animal can make use of the inorganic phosphate liberated from phytate phosphorus using the

enzyme. The phytase formulation of the invention has an improved thermostability and can therefore remain stable during long-term storage and can withstand feed processing methods such as extrusion, expansion and pelleting. The present sequence represents an initial consensus phytase sequence determined using the program PRETTY from mature phytase sequences from a variety of fungi (AAY69544-Y69546, AAY69548-Y69556, AAY69564) and the Basidiomycetes phytase consensus AAY69563. This was used as the basis for the consensus phytase designated phytase-10 (AAY69566).

XX Sequence 424 AA;

Query Match 88.9%; Score 2064.5; DB 21; Length 424;
 Best Local Similarity 94.3%; Pred. No. 2.5e-193;
 Matches 414; Conservative 0; Mismatches 8; Indels 17; Gaps 12;
 QY 1 NSHSCDITVD-GYCPETISHLWQYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARYP 59
 DB 1 nshscdtvdggyqcpelshlwqyspffsladesaispdvp-gcrvtfvqvlsrhgaryp 59
 QY 60 TSSSKKYSALIERIQKNATFKGYAFLKTYNTLGGADDLTPFGENOMVNSGKIFVRRYK 119
 DB 60 tssskkysali-aiqknatfkkyafikcynyclgaddltfpgeqmwvnsqkifvrryk 118
 QY 120 ALARNIVPFRASGSDRVIASAEKFIQFQSAKLADPAHOASPVINVIIPESGSGYNNTLD 179
 DB 119 alar-ivpfrasgsdrviasaekfiqfqsakladpa-qaspvinvlipeg-gynntld 175
 QY 180 HGLCTAFEDSTLGGDAEANTFAVPAPIRARLEALPCVNLTDDEVDVNLMDMCPFDVART 239
 DB 176 hglctafepseldgveantfappirarlealpcvnltdedvnlmdmcpfdvart 234
 QY 240 SDATOLSPFCDLFTADEWOYDLOSLKYYGYGAGNPLGPAQGGVFNELIARLTHSPVQDH 299
 DB 235 sdatslspfcldft-hewdydylslkygygagnpplgpaggvgfnellarlthspvqdh 293
 QY 300 TSTNHTLDSNPATPLNATLYADFSDHNTWVSIFPGLGLYNGTKPLSTTSVE-SIETDGY 358
 DB 294 tstnhtldsnpatplnatlyadfsdntwvsiiffalglngtkplsttsve-sietdgy 352
 QY 359 AASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRVLVNDVRVPLHGCYDKLGRCKLDD 418
 DB 353 aaswtvpfaarayvemmqce-----egekeplvrvlvndrvrplhgcgvdkgcklgrck-dd 405
 QY 419 FVEGLSFARSGGNWAECEFA 437
 DB 406 fveglstfarsggnweecfa 424

RESULT 14

AAB20532
 ID AAB20532 standard; Protein; 467 AA.

XX
 AC AAB20532;
 XX

XX 05-DEC-2000 (first entry)

Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.

Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 temperature stability; pH profile; temperature profile; reaction rate;
 specific activity; substrate specificity; substrate cleavage pattern;
 substrate binding; position specificity; phytate degradation rate;
 food; feed; phytate; manure.

XX Synthetic.
 OS
 XX
 XX WC200043503-A1.
 XX 27-JUL-2000.
 XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Lehmann M;
 XX
 XX WPI; 2000-491161/43.
 DR N-PSDB; AAA73291.
 XX
 PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX
 XX Disclosure; Fig 23a-c; 240pp; English.
 PS
 XX
 CC The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.
 XX
 SQ Sequence 467 AA;

Query Match 88.8%; Score 2061; DB 21; Length 467;
 Best Local Similarity 91.1%; Pred. No. 6.5e-193;
 Matches 408; Conservative 7; Mismatches 15; Indels 18; Gaps 10;
 QY 1 NSHSCDTPD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARY 58
 DB 27 nshscdtdvggygcfeishlwgtyspysfsladesaispdvpgkdrvtfvqlsthrgary 86
 QY 59 PTSSKSKYSALIERIQKNAT-FKGYAFKTYNTLGGADDLTPFGENQMVNSGKIFYRR 117
 DB 87 ptssaskaysalialcqnatafkgyafktyntylgaddltpfgenqmvnsgikfyrr 146
 QY 118 YKALARNIVPVRASGSDRVIASAEKFIQGSQAKLADPA---HOASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaekfiegfsqakladpgsqhqsqspvinnvipegsgy 206
 QY 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafedstlgdgveanftalfapairleadlpgvntldedvvyldmcpf 266
 QY 234 DTVARTSDATQSPCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatelspcaiftdhewiqydyloslgkygygagngnplgpaqgvgfaneliar 326
 QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFALGLYNGTKPLSTTSV 350
 DB 327 lthspvqdhstntnhtlidsnpatfplnactlyadfsdntmvsifalglyngtkplsttsv 386
 QY 351 RSI-BTDGYASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDRVVPLHGCVD 409
 DB 387 esietdgyaswtvtpfaarayvemmqcqa-----ekeplvrlvndrvvplhgcavd 439
 QY 410 KLGRCKLDDFVEGLSFARSNGNWAECFA 437
 DB 440 klgrckrdddfveglsfarsngnwaecfa 467

RESULT 15
 AAY43169

ID AAY43169 standard; Protein; 467 AA.
 XX
 AC AAY43169;
 XX
 DT 06-JAN-2000 (first entry)
 XX
 DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.
 XX
 KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 KW consensus sequence.
 XX
 OS Synthetic.
 XX
 XX WO9948380-A1
 PN PD 30-SEP-1999.
 XX
 XX 22-MAR-1999; 99WO-DK00154.
 PF
 XX 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 PR 22-JAN-1999; 99DK-0000093.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 XX Petersen S;
 PI
 DR WPI; 1999-591030/50.
 DR N-PSDB; AAZ31520.
 XX
 PT Preparing animal feed using a thermostable phytase
 XX
 PS Example 3; Fig 9; 7lpp; English.
 XX
 CC This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.
 XX
 SQ Sequence 467 AA;

Query Match 88.1%; Score 2044; DB 20; Length 467;
 Best Local Similarity 90.4%; Pred. No. 3e-191;
 Matches 405; Conservative 7; Mismatches 18; Indels 18; Gaps 10;
 QY 1 NSHSCDTPD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGCRVTFVQVLSRHGARY 58
 DB 27 nshscdtdvggygcfeishlwgtyspysfsladesaispdvpgkdrvtfvqlsthrgary 86
 QY 59 PTSSKSKYSALIERIQKNAT-FKGYAFKTYNTLGGADDLTPFGENQMVNSGKIFYRR 117
 DB 87 ptssaskaysalialcqnatafkgyafktyntylgaddltpfgenqmvnsgikfyrr 146
 QY 118 YKALARNIVPVRASGSDRVIASAEKFIQGSQAKLADPA---HOASPVINVIPEGSGY 174
 DB 147 ykalarkivpfirasgsdrviasaekfiegfsqakladpgsqhqsqspvinnvipegsgy 206
 QY 175 NNTLDHGLCTAFEDSTLGDGAENFTAVFAPPFIRARLEA-LPGVNLTDDEVVNLMDMCPF 233
 DB 207 nntldhglctafedstlgdgveanftalfapairleadlpgvntldedvvyldmcpf 266
 QY 234 DTVARTSDATQSPCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
 DB 267 dtvartsdatelspcaiftdhewiqydyloslgkygygagngnplgpaqgvgfaneliar 326

Sat Oct 27 15:25:41 2001

QY 291 LTHSPVODHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFFALGLYNGTKPLSTTSV 350
Db 327 lthspvqdhstnhtlidsnpatfplnatlyadfsdntmisiffalglyngtkplsttsv 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMQCEAGGGGEGEKEPELYRVLVNDRVVPLHGCQVD 409
Db 387 esieetdgysaswtvpfaarayvemmqcqa-----ekeplrvivndrvvpplhgcavd 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 klgrckrddfvglsfarsggnwaecefa 467

Search completed: October 26, 2001, 16:40:04
Job time: 4955 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:45 ; Search time 32.57 Seconds
(without alignments)
276.266 Million cell updates/sec

Title: US-09-488-265-27
Perfect score: 2321
Sequence: 1 NSHSCDVTVDGYOCPEISHLW.....DFVEGLSFARGGNWAECEFA 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	87.0	441	4	US-09-121-425-1
2	1902	81.9	467	4	US-09-121-425-2
3	1688	72.7	467	1	US-07-923-724-8
4	1688	72.7	467	2	US-08-609-426A-8
5	1688	72.7	467	2	US-08-374-652C-2
6	1684	72.6	467	1	US-08-151-574-32
7	1684	72.6	467	1	US-08-146-424-20
8	1684	72.6	467	1	US-08-693-709-2
9	1684	72.6	467	2	US-08-419-448-32
10	1684	72.6	467	2	US-08-819-825-3
11	1684	72.6	467	4	US-09-163-642-3
12	1672	72.0	443	4	US-09-155-855-1
13	1672	72.0	443	4	US-09-155-855-3
14	1667	71.8	443	4	US-09-155-855-2
15	1322	57.0	475	4	US-08-819-825-2
16	1322	57.0	475	4	US-09-163-642-2
17	837.5	36.1	443	3	US-08-993-359-30
18	812.5	35.0	453	3	US-08-993-359-22
19	804.5	34.7	439	3	US-08-993-359-24
20	804.5	34.7	439	3	US-08-221-654-2
21	793	34.2	442	3	US-08-989-358A-2
22	791.5	34.1	442	3	US-08-993-359-28
23	343.5	14.8	468	1	US-07-627-539G-7
24	343.5	14.8	468	1	US-07-627-539G-2
25	307.5	13.2	479	1	US-07-923-724-2
26	307.5	13.2	479	2	US-08-609-426A-2
27	307.5	13.2	479	2	US-08-609-426A-3

28	307.5	13.2	479	2	US-08-374-652C-4	Sequence 4, Appli
29	169	7.3	92	3	US-08-993-359-32	Sequence 32, Appli
30	166	7.2	449	3	US-08-680-506-7	Sequence 7, Appli
31	115	5.0	318	3	US-08-680-506-3	Sequence 32, Appli
32	104.5	4.5	24	2	US-08-374-652C-32	Sequence 3, Appli
33	101.5	4.4	297	2	US-09-027-013-3	Sequence 3, Appli
34	101.5	4.4	297	3	US-09-244-233-3	Sequence 3, Appli
35	100.5	4.3	113	1	US-08-241-853-8	Sequence 8, Appli
36	100.5	4.3	113	2	US-08-850-917-8	Sequence 8, Appli
37	98	4.2	20	1	US-07-923-724-43	Sequence 43, Appli
38	98	4.2	20	2	US-08-609-426A-43	Sequence 43, Appli
39	93.5	4.0	750	6	5457037-3	Patent No. 5457037
40	93.5	4.0	751	6	5457037-5	Patent No. 5457037
41	90.5	3.9	113	1	US-08-241-853-10	Sequence 10, Appli
42	90.5	3.9	113	2	US-08-850-917-10	Sequence 10, Appli
43	88	3.8	331	4	US-08-849-751-4	Sequence 4, Appli
44	87.5	3.8	499	2	US-09-032-315-3	Sequence 3, Appli
45	87.5	3.8	499	2	US-08-993-318A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-1

Query Match	87.0%	Score 2020;	DB 4;	Length 441;
Best Local Similarity	89.3%	Pred. No. 6e-208;		
Mismatches	400;	Conservative	10;	Mismatches 20;
Indels	18;	Gaps	10;	
QY	1	NHSCDVTVD-GYQC-PEISHLWGQYSPFFSLADESAISPDVPGRCRVTFVQVLSRHGARY	58	
Db	1	NHSCDVTVDGQYQCPEISHLWGQYSPFFSLEDESAISPDVPGRCRVTFVQVLSRHGARY	60	
QY	59	PTSSKSKYSALIERIGKNAT-FKGKVAFLKTYNTYLGADLTPFGENQMVNSGKIFYR	117	
Db	61	PTSSKSKYSALIERIGKNAT-FKGKVAFLKTYNTYLGADLTPFGENQMVNSGKIFYR	120	
QY	118	YKALARNIVFVRASGSDRVIAEAKEFIEGFSQAKLADPA---HQASPTVINIIPESGSY	174	
Db	121	YKALARKIVFIFRASGSDRVIAEAKEFIEGFSQAKLADPA---HQASPTVINIIPESGSY	180	
QY	175	NNTLDHGLCTAFEDSTLGDAAEANTAVEAPPTRARLEA-LPGVNLTDDEVNLMDCPF	233	
Db	181	NNTLDHGLCTAFEDSTLGDAAEANTAVEAPPTRARLEA-LPGVNLTDDEVNLMDCPF	240	
QY	234	DTVARTSDATQLSPPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR	290	
Db	241	ETVARTSDATELSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR	300	
QY	291	LTHSPVQDHTSTNHTLDSNPATFPPLNATLYADFSHNTMWSIFFALGLYNGTKPLSTTSV	350	

```

Db 301 LTRSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNSMISIPFALGLVNGTAPLSTTSV 360
Qy 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDVRVPLHOCGYD 409
Db 361 ESIETDGYASWTVPFARAYVEMMOCEA-----EKEPLVRLVNDVRVPLHGCAYD 413
Qy 410 KLGCKLDDFVEGLSFARSGGNAECFA 437
Db 414 KLGCKRDDFVEGLSFARSGGNAECFA 441

RESULT 2
US-09-121-425-2
; Sequence 2, Application us/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: sequence
US-09-121-425-2

Query Match 81.9%; Score 1902; DB 4; Length 467;
Best Local Similarity 81.8%; Pred. No. 2.9e-195;
Matches 382; Conservative 10; Mismatches 19; Indels 56; Gaps 11;

Qy 1 NSHSCDVTVD-GYQC-PEISHLWQGYSPFFSLADESAISPDVPKGRVTFVQVLSRHGARY 58
Db 27 NSHSCDVTVDGQYQCFPEISHLWQGYSPFFSLADESAISPDVPKGRVTFVQVLSRHGARY 86
Qy 59 PTSSKSKYSALIEROKNATEFGKYAFKLTNYTLGADDLTPFGENQVNSGKIFRYR 118
Db 87 PTSSKSKYSA-----TNYTLGADDLTPFGENQVNSGKIFRYR 127
Qy 119 KALARNIVPFRASGSDRYASAEKFIQFQSAKLADPA---HQASPVIN----- 165
Db 128 KALARKIVPFRASGSDRYASAEKFIQFQSAKLADPGSQPHQASPVLDLIEATQKNAT 187
Qy 166 -----VTIPEGSGYNNLIDHGLCTAFEDSTLGDAAEANTFAVAPPFIRARLEA-L 214
Db 188 AFGKYAFKLTNYIPEGSGYNNLIDHGLCTAFEDSTLGDAAEANTFAVAPPFIRARLEADL 247
Qy 215 PGVNLTDVDVNLMDMCPDFTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYCA 272
Db 248 PGVNLTDVDVNLMDMCPETVARTSDATQSPFCDLFTADEWQYDYLQSLKYGYCA 307
Qy 273 GNPLGPAQGVGF-NELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFSDHNTWVS 331
Db 308 GNPLGPAQGVGFANELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFSDHNSMIS 367
Qy 332 IFPAGLYNGTKPLSTSVESI-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEP 390
Db 368 IFPAGLYNGTAPLSTSVESIETDGYASWTVPFARAYVEMMOCEA-----EKEP 420
Qy 391 LVRVLNDRVPLHGGGVCKLGRCKLDDFVEGLSFARSGGNAECFA 437
Db 421 LVRVLNDRVPLHGGCAVDKLGCKRDDFVEGLSFARSGGNAECFA 467

```

```

US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimdala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8

Query Match 72.7%; Score 1688; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 2.6e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

Qy 1 NSHSCDVTVD-GYQC-PEISHLWQGYSPFFSLADESAISPDVPKGRVTFVQVLSRHGARY 58
Db 27 NOSHCDVTVDQYQCFSETSHLWQYAPFFSLANESALSPDPAGCRVTFVQVLSRHGARY 86
Qy 59 PTSSKSKYSALIEROKNATEFGKYAFKLTNYTLGADDLTPFGENQVNSGKIFRYR 117
Db 87 PTESKGGYSALIEEQNVTFDQKYAFKLTNYSLGADDLTPFGQELVNSGKIFYOR 146
Qy 118 YKALARNIVPFRASGSDRYASAEKFIQFQSAKLADP---HQASPVINVIPEGSGY 174
Db 147 YESLTRNIIPFIRSSGSRVIASGEKIEGFOSTKLKDPRAQPGQSGPKIDVVISEASS 206

```



```

QY 175 NNTLDHGLCTAFEDSTLGDAAEFANFTAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233
Db 207 NNTLDPGCTCTVFEDSELADTVEANFTAFAPSIQRLENLDSGVTLTDTETVYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTDEWHIDYLOSLKYYGAGNPLGPTQGVGVANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHONTWYSIFALGLYNGTLPSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSHONGIISILFALGLYNGTLPSTTSV 386
QY 351 ESI-ETDGYAASWTVPFAARYVEMMOCEAGGGGEGEKEPLVRVLNDRVPLHCGVD 409
Db 387 ENITQDTDFSSAWTVPFASRLYVEMMQCA-----EQEPLVRVLNDRVPLHCGPID 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECSA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 4
US-08-609-426A-8
; Sequence 8, Application US/08609426A
; Patent No. 5830733
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; APPLICANT: Houston, Christine S.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; IN TRICHODERMA
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,426A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,724
; FILING DATE: 31-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Grant E.
; REGISTRATION NUMBER: P-41,264
; REFERENCE/DOCKET NUMBER: 1050.0080001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600

```

```

; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-609-426A-8

Query Match
Best Local Similarity 72.7%; Score 1688; DB 2; Length 467;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NNSHCDTVD-QYQC-PEJSHLWGOYSPFFSLADESAISPDVPKGCRTVTFVOLSRRHGARY 58
Db 27 NOSTCDTVDQGYQCFSETHLWGOYAPFFSLANESAISSPDVPAGCRVTFVOLSRRHGARY 86
QY 59 PTSSKSKYSALIERIQKN-ATFKGYAFLKTYNTLCADDLTPFGENOMVNSGIKEYRR 117
Db 87 PTESKSKYSALIEIQONVTTFDGYAFLKTYNSLGADDLTFEGEQLVNSGIKEYRR 146
QY 118 YKALARNIVPFRASGSDRVITASAEKFTIEGQSALADP---AHQASPVINVIIEGSGY 174
Db 147 YESLTRNIIPFIRSSGSRVITASAEKFTIEGQSALADP---AHQASPVINVIIEGSGY 206
QY 175 NNTLDHGLCTAFEDSTLGDAAEFANFTAVFAPPFIRARLE-ALPGVNLTDDEVDVNLMDMCPF 233
Db 207 NNTLDPGCTCTVFEDSELADTVEANFTAFAPSIQRLENLDSGVTLTDTETVYLMDCSF 266
QY 234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLOSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTDEWHIDYLOSLKYYGAGNPLGPTQGVGVANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHONTWYSIFALGLYNGTLPSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSHONGIISILFALGLYNGTLPSTTSV 386
QY 351 ESI-ETDGYAASWTVPFAARYVEMMOCEAGGGGEGEKEPLVRVLNDRVPLHCGVD 409
Db 387 ENITQDTDFSSAWTVPFASRLYVEMMQCA-----EQEPLVRVLNDRVPLHCGPID 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECSA 437
Db 440 ALGRCTRDSFVRGLSFARSGGDWAECSA 467

```

```

RESULT 5
US-08-374-652C-2
; Sequence 2, Application US/08374652C
; Patent No. 5834286
; GENERAL INFORMATION:
; APPLICANT: NEVALAINEN, HELENA K.M.
; APPLICANT: PALOHEIMO, MARJA T.
; APPLICANT: FAGERSTROM, RICHARD B.
; APPLICANT: MIETTINEN-OINONEN, ARJA S.
; APPLICANT: TURUNEN, MARJA K.
; APPLICANT: RAMBOSEK, JOHN A.
; APPLICANT: PIDDINGTON, CHRISTOPHER S.
; APPLICANT: HOUSTON, CHRISTINE S.
; APPLICANT: CANTRELL, MICHAEL A.
; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
; VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
; ENZYMES IN DESIRED RATIOS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:

```

us-09-488-265-27.ra1

Sat Oct 27 15:25:42 2001

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-374-652C-2

```

```

Query Match 72.7%; Score 1688; DB 2; Length 467;
Best Local Similarity 73.9%; Pred. No. 2.6e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NOSTCDTVDGQYCFSETSHLWGOYAFPPFLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIOTKN-ATFKGKYAFKTYNTLIGADDLTPFGENQMVNSGKIFYRR 117
Db 87 PTESKSKYSALIEIQNVTFDCKYAFKTYNTSLGADDLTPFGEQELVNSGKIFYOR 146
QY 118 YKALARNIVPVRASGSDRVASAEKIEFGQSAAKLADP---AHOASPVINVIPEGSGY 174
Db 147 YESLTRNIIPFIRSSGSRVIAASGKIEFGQSAAKLADP---AHOASPVINVIPEGSGY 206
QY 175 NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPFIRARLE-ALPGVNLTDVVDVNLMDMCPF 233
Db 207 NNTLDPGTCTVFDESELADTVFANFTATFAPSIRQLENDLSGVTLTDEVTYLMDCSF 266
QY 234 DTVARTSDATQSPFCDLFTADEW-OYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTHDEWIHYDYLQSLKYYGHGAGNPLGPTQGVGYANEIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSDHNTMVSIFPALGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASVTVPFAARVYVEMMQCEAGGGGGEKEPLRVRLVNDRVVPLHGCGVD 409
Db 387 ENITQDGFSSANTVFPASRLYVEMMQCA-----EQEPLRVRLVNDRVVPLHGCPID 439
QY 410 KLGRCKLDDFVEGLSFARSGNNAECFA 437
Db 440 ALGRCTRDSFVRGLSEFARSGDWAECFA 467

```

RESULT 6

US-08-151-574-32

; Sequence 32, Application US/08151574

```

; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-151-574-32

```

```

Query Match 72.6%; Score 1684; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGOYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NOSTCDTVDGQYCFSETSHLWGOYAFPPFLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIOTKN-ATFKGKYAFKTYNTLIGADDLTPFGENQMVNSGKIFYRR 117
Db 87 PTDSKSKYSALIEIQNVTFDCKYAFKTYNTSLGADDLTPFGEQELVNSGKIFYOR 146
QY 118 YKALARNIVPVRASGSDRVASAEKIEFGQSAAKLADP---AHOASPVINVIPEGSGY 174
Db 147 YESLTRNIIPFIRSSGSRVIAASGKIEFGQSAAKLADP---AHOASPVINVIPEGSGY 206
QY 175 NNTLDHGLCTAFEDSTLGDAAENFTAVFAPPFIRARLE-ALPGVNLTDVVDVNLMDMCPF 233
Db 207 NNTLDPGTCTVFDESELADTVFANFTATFAPSIRQLENDLSGVTLTDEVTYLMDCSF 266
QY 234 DTVARTSDATQSPFCDLFTADEW-OYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTHDEWIHYDYLQSLKYYGHGAGNPLGPTQGVGYANEIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTMVSIFPALGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATPLNATLYADFSDHNTMVSIFPALGLYNGTKPLSTTV 386

```

```

QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGCGGEGEKEPLVRLVNDVVRVPLHGCQVD 409
Db 387 ENITQDGFSSAWTVPFASRLYVEMMOCEA-----EQEPLVRLVNDVVRVPLHGCQVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDSEFVRGLSFARSGGNWAECEFA 467

RESULT 7
US-08-146-424-20
; Sequence 20, Application US/08146424
; Patent No. 5593963
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,424
; FILING DATE: 02-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KENNEDY, BILL
; REGISTRATION NUMBER: 33,407
; REFERENCE/DOCKET NUMBER: 44615-20011.24
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-146-424-20

Query Match 72.6%; Score 1684; DB 1; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVTVD-GYQC-PEISHLWGQYSPFSLADESAISPDVPGKGVTFYQVLSRHGARY 58
Db 27 NOSSCDTVDQGYQCFSETSHLWGQYAPFSLANESVISPVPACRVTFYQVLSRHGARY 86
QY 59 PTSSKSKYSALIEIQKNA-TFFGKYAFKTYNTTGADLTTPFGQELVNSGKIFQYR 117
Db 87 PTDSKGYKYSALIEIQKNA-TFFGKYAFKTYNTTGADLTTPFGQELVNSGKIFQYR 146
QY 118 YKALARNIVPVRAGSGSDRVIASAEKTEGFQSAKLADP---AHOASPVNVIPEGSGY 174
Db 147 YESLRNIVPVRAGSGSDRVIASAEKTEGFQSAKLADP---AHOASPVNVIPEGSGY 206
QY 175 NNTLDHGLCTAFEDSTLGDGDAENFTAVFAPP IRRARLE-ALPGVNLTDVWVNLMDMCPF 233

```

```

Db 207 NNTLDPGCTCTVPEDSELADTVEANFTATVPSTIQRLENLDSGVLTDEVTYLMDMCSF 266
QY 234 DTVARTSDATOLSPFCDLTADW-QYDYLQSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLTADW-QYDYLQSL-KYYGYGAGNPLGPAQGVGF-NELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFFALGYNKTPKLSSTVS 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSDHNTMVSIFFFALGYNKTPKLSSTVS 386
QY 351 ESI-ETDGYAASWTVPFAARAYVEMMOCEAGGCGGEGEKEPLVRLVNDVVRVPLHGCQVD 409
Db 387 ENITQDGFSSAWTVPFASRLYVEMMOCEA-----EQEPLVRLVNDVVRVPLHGCQVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437
Db 440 ALGRCTRDSEFVRGLSFARSGGNWAECEFA 467

RESULT 8
US-08-693-709-2
; Sequence 2, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OOLJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1...23
; OTHER INFORMATION:

```

US-08-693-709-2

Query Match 72.6%; Score 1684; DB 1; Length 467;

Best Local Similarity 73.9%; Pred. No. 6.9e-172; Indels 18; Gaps 10;

Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

1 NNSHCDTVD-GYOC-PEISHLMGQYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58

27 NQSSCDTVDQGYQCFSETSHLMGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86

59 PTSSKSKYSALIERIOKNA-TFKGKYAFKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117

87 PTDSKGRKYSALIEIOQATTFDGYAFKTYNTLTGADDLTPFGEQELVNSGKIFYYR 146

118 YKALARNIVPVRASGSDRVIASAEKIEGFQSAKLADP---AHQASPVINVIPEGSGY 174

147 YESLTRNIVPFISSGSRVIRASGKIEGFQSTKLDPRAGQCSGPKIDVIVISEASS 206

175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPIRABLE-ALPGVNLTDVNVLMDCMCF 233

207 NNTLDPGCTVFEDSELADTVEANFTATFVPSIRQLENDLSGVTLDTEVTYLMDCSF 266

234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290

267 DTISTSTVDTKLSPCDLFTHDEWINYDYLQSLKYYGAGNPLGPAQGVGF-NELIAR 326

350 LTHSPVQDHTSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTKPLSTTSV 350

327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTKPLSTTSV 386

351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGEKEPLRVLVNDRVPLHGCQVD 409

387 ENITQDTGFSASWTVPFASRLYVEMMQCOA-----EQEPLRVLVNDRVPLHGCQVD 439

410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437

440 ALGRCTRDSFVRGLSFARSGGNWAECEFA 467

RESULT 9

US-08-419-448-32

Sequence 32, Application US/08419448

Patent No. 5863533

GENERAL INFORMATION:

APPLICANT: Robert F.M. Van Gorcom

APPLICANT: Willem Van Hartingsveldt

APPLICANT: Petrus A. Van Paridon

APPLICANT: Annemarie E. Veenstra

APPLICANT: Rudolf G.M. Luttin

TITLE OF INVENTION: Cloning and Expression of Microbial

TITLE OF INVENTION: Phytase

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Ave. N.W., Suite 5500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

SOFTWARE:

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/419,448

FILING DATE: 10-APR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murashige, Kate H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 24615-20026.10

TELEPHONE: 202-887-1500

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-419-448-32

Query Match 72.6%; Score 1684; DB 2; Length 467;

Best Local Similarity 73.9%; Pred. No. 6.9e-172;

Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

1 NNSHCDTVD-GYOC-PEISHLMGQYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58

27 NQSSCDTVDQGYQCFSETSHLMGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86

59 PTSSKSKYSALIERIOKNA-TFKGKYAFKTYNTLTGADDLTPFGENQMVNSGKIFYYR 117

87 PTDSKGRKYSALIEIOQATTFDGYAFKTYNTLTGADDLTPFGEQELVNSGKIFYYR 146

118 YKALARNIVPVRASGSDRVIASAEKIEGFQSAKLADP---AHQASPVINVIPEGSGY 174

147 YESLTRNIVPFISSGSRVIRASGKIEGFQSTKLDPRAGQCSGPKIDVIVISEASS 206

175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPIRABLE-ALPGVNLTDVNVLMDCMCF 233

207 NNTLDPGCTVFEDSELADTVEANFTATFVPSIRQLENDLSGVTLDTEVTYLMDCSF 266

234 DTVARTSDATQSLSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290

267 DTISTSTVDTKLSPCDLFTHDEWINYDYLQSLKYYGAGNPLGPTQGVGYANELIAR 326

291 LTHSPVQDHTSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTKPLSTTSV 350

327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAALGLYNGTKPLSTTSV 386

351 EST-ETDGYAASWTVPFAARAYVEMMQCEAGGGEKEPLRVLVNDRVPLHGCQVD 409

387 ENITQDTGFSASWTVPFASRLYVEMMQCOA-----EQEPLRVLVNDRVPLHGCQVD 439

410 KLGRCKLDDFVEGLSFARSGGNWAECEFA 437

440 ALGRCTRDSFVRGLSFARSGGNWAECEFA 467

RESULT 10

US-08-819-825-3

Sequence 3, Application US/08819825

Patent No. 5866118

GENERAL INFORMATION:

APPLICANT: Berk, Randy M.

APPLICANT: Ray, Michael W.

APPLICANT: Klotz, Alan V.

TITLE OF INVENTION: Polypeptides Having Phytase Activity

TITLE OF INVENTION: And Nucleic Acids Encoding Same

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for Windows Version 2.0

COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-819-825-3

Query Match 72.6%; Score 1684; DB 2; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PEISHLWGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSCDVTDOGYQCFSESHLWGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIQKNA-TFKGYAFLKTYNTLGADDLTPFGENQMVNSGKIFYRR 117
Db 87 PTDSKGGKYSALIEEQONATTFDGYAFKTYNTSLGADDLTPFGQELVNSGKIFYQR 146
QY 118 YKALARNIVPVRASGSDRVIAAEKIEGFSQAKLADP---AHOASPVNVIIPESGY 174
Db 147 YESLTRNIVPIRSSGSRVIAAGKIEGFSQAKLADP---AHOASPVNVIIPESGY 206
QY 175 NNTLDHGLCTAFEDSLTGLDAAENFTAVFAPPRIARLE-ALPGVNLTDDEVNLMDCPF 233
Db 207 NNTLDPGTCTVFEDSELATVEANFTATFVPSIRQLENLDSGLTLDTEVYLMDCSF 266
QY 234 DTIVARTSDATQSLPFCDLFTADEW-QYDYLOSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTHDEWINYDYLQSLKKYGHGAGNPLGPTQGVYANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDRVVPLHGCVD 409
Db 387 ENITQTDGFSSTAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDRVVPLHGCVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437
Db 440 ALGRCRTDSFVRGLSFARSGGDWAECA 467

RESULT 12
US-09-163-642-3
Sequence 3, Application US/09163642
Patent No. 6221644
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Ray, Michael W.
APPLICANT: Klotz, Alan V.
TITLE OF INVENTION: Polypeptides Having Phytase Activity
TITLE OF INVENTION: And Nucleic Acids Encoding Same
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6221644 No. 6221644disk of No. 6221644th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York

COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-163-642-3

Query Match 72.6%; Score 1684; DB 4; Length 467;
Best Local Similarity 73.9%; Pred. No. 6.9e-172;
Matches 331; Conservative 40; Mismatches 59; Indels 18; Gaps 10;

QY 1 NSHSCDVID-GYQC-PEISHLWGQYSPFFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
Db 27 NQSCDVTDOGYQCFSESHLWGQYAPFFSLANESVISPVPAGCRVTFVQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIQKNA-TFKGYAFLKTYNTLGADDLTPFGENQMVNSGKIFYRR 117
Db 87 PTDSKGGKYSALIEEQONATTFDGYAFKTYNTSLGADDLTPFGQELVNSGKIFYQR 146
QY 118 YKALARNIVPVRASGSDRVIAAEKIEGFSQAKLADP---AHOASPVNVIIPESGY 174
Db 147 YESLTRNIVPIRSSGSRVIAAGKIEGFSQAKLADP---AHOASPVNVIIPESGY 206
QY 175 NNTLDHGLCTAFEDSLTGLDAAENFTAVFAPPRIARLE-ALPGVNLTDDEVNLMDCPF 233
Db 207 NNTLDPGTCTVFEDSELATVEANFTATFVPSIRQLENLDSGLTLDTEVYLMDCSF 266
QY 234 DTIVARTSDATQSLPFCDLFTADEW-QYDYLOSL-KYYGYGAGNPLGPAQGVGF-NELIAR 290
Db 267 DTISTSTVDTKLSPFCDLFTHDEWINYDYLQSLKKYGHGAGNPLGPTQGVYANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTSV 350
Db 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNTMVSIFFAFGLYNGTKPLSTTV 386
QY 351 EST-ETDGYAASWTVPFAARAYVEMMOCEAGGGGEGEKEPLVRLVNDRVVPLHGCVD 409
Db 387 ENITQTDGFSSTAWTVPFASRLYVEMMOCA-----EQEPLVRLVNDRVVPLHGCVD 439
QY 410 KLGRCKLDDFVEGLSFARSGGNWAECA 437
Db 440 ALGRCRTDSFVRGLSFARSGGDWAECA 467

RESULT 12
US-09-155-855-1
Sequence 1, Application US/09155855
Patent No. 6139902
GENERAL INFORMATION:
APPLICANT: KONDO, Hidemasa

```

; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-1

```

```

Query Match 72.0%; Score 1672; DB 4; Length 443;
Best Local Similarity 72.8%; Pred. No. 1.2e-170;
Matches 326; Conservative 46; Mismatches 58; Indels 18; Gaps 10;

QY 1 NSHSCDTVD-GYQC-PEISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
DB 3 NOSTCDTVDGQYCFSETSHLWGOYAPFFSLANKSAISPDVPAGCHVTFQAQVLSRHGARY 62
QY 59 PTSSKSKYSALIERIQKNA-TFKGKYAFLKTYNTLTGADDLTPFGENQMVNSGKIFYRR 117
DB 63 PTDSKGGKYSALIEIQONATTFECKYAFKTYNYSLGADDLTPFGEQELVNSGVKFYQR 122
QY 118 YKALARNIVPFVRASGSDRVIASAEKFTFEGQSALADP---AHQASPVINVIIEGSGY 174
DB 123 YESLTRNIVPFRSSGSSRVIASGNKFIEGFQSTKLKDPRAQGOSSPKIDVISEASTS 182
QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPFIRARLE-ALPGVNLTDDEYVNLMDMCPF 233
DB 183 NNTLDPGCTCTVEDESELADDEANFTATVPFVPSIRQLENDLSGVSITDEVTYLMDCSF 242
QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 243 DTISTSTVDTKLSPFCDLFTHEEWINYDLSLNKYHGGAGNPLGPTQGVGYANELIAR 302
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPALGLYNGTKPLSTTSV 350
DB 303 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSDNGIISILFALGLYNGTKPLSSTA 362
QY 351 EST-ETDGYAASVTVPFAARAYVEMMQCEAGGGGEGEKEPLRVRLVNDVRVPLHGCGVD 409
DB 363 ENITQIDGFSANTVPFASRMVEMMQCS-----EQEPLRVRLVNDVRVPLHGCPVD 415
QY 410 KLGRCLDDFVEGLSFARSGNWAECFA 437
DB 416 ALGRCTRDSEVFKLSFARSGDNGECFA 443

```

```

RESULT 13
US-09-155-855-3
; Sequence 3, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175

```

```

; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-3

Query Match 72.0%; Score 1672; DB 4; Length 467;
Best Local Similarity 72.8%; Pred. No. 1.3e-170;
Matches 326; Conservative 46; Mismatches 58; Indels 18; Gaps 10;

QY 1 NSHSCDTVD-GYQC-PEISHLWGOYSPFSLADESAISPDVPGKCRVTFVQVLSRHGARY 58
DB 27 NOSTCDTVDGQYCFSETSHLWGOYAPFFSLANKSAISPDVPAGCHVTFQAQVLSRHGARY 86
QY 59 PTSSKSKYSALIERIQKNA-TFKGKYAFLKTYNTLTGADDLTPFGENQMVNSGKIFYRR 117
DB 87 PTDSKGGKYSALIEIQONATTFECKYAFKTYNYSLGADDLTPFGEQELVNSGVKFYQR 146
QY 118 YKALARNIVPFVRASGSDRVIASAEKFTFEGQSALADP---AHQASPVINVIIEGSGY 174
DB 147 YESLTRNIVPFRSSGSSRVIASGNKFIEGFQSTKLKDPRAQGOSSPKIDVISEASTS 206
QY 175 NNTLDHGLCTAFEDSTLGDAAEANTAVFAPPFIRARLE-ALPGVNLTDDEYVNLMDMCPF 233
DB 207 NNTLDPGCTCTVEDESELADDEANFTATVPFVPSIRQLENDLSGVSITDEVTYLMDCSF 266
QY 234 DTVARTSDATQSPFCDLFTADEW-QYDYLQSL-KYGYGAGNPLGPAQGVGF-NELIAR 290
DB 267 DTISTSTVDTKLSPFCDLFTHEEWINYDLSLNKYHGGAGNPLGPTQGVGYANELIAR 326
QY 291 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPALGLYNGTKPLSTTSV 350
DB 327 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSDNGIISILFALGLYNGTKPLSSTA 386
QY 351 EST-ETDGYAASVTVPFAARAYVEMMQCEAGGGGEGEKEPLRVRLVNDVRVPLHGCGVD 409
DB 387 ENITQIDGFSANTVPFASRMVEMMQCS-----EQEPLRVRLVNDVRVPLHGCPVD 439
QY 410 KLGRCLDDFVEGLSFARSGNWAECFA 437
DB 440 ALGRCTRDSEVFKLSFARSGDNGECFA 467

```

```

RESULT 14
US-09-155-855-2
; Sequence 2, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JF97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger

```


Sat Oct 27 15:25:42 2001

us-09-488-265-27.ra1

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 18:33:00 ; Search time 2563.44 Seconds
(without alignments)
5177.342 Million cell updates/sec

Title: US-09-488-265-28
Perfect score: 1404
Sequence: 1 atggcgctgttcgtctgct.....ggcgctgaatgttcgtctaa 1404

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_est4:*
 - 5: gb_est5:*
 - 6: gb_est6:*
 - 7: gb_est7:*
 - 8: gb_est8:*
 - 9: gb_est9:*
 - 10: gb_est10:*
 - 11: gb_est11:*
 - 12: gb_est12:*
 - 13: gb_est13:*
 - 14: gb_est14:*
 - 15: gb_est15:*
 - 16: gb_est16:*
 - 17: gb_est17:*
 - 18: gb_est18:*
 - 19: gb_est19:*
 - 20: gb_est20:*
 - 21: gb_est21:*
 - 22: gb_est22:*
 - 23: gb_est23:*
 - 24: gb_est24:*
 - 25: gb_est25:*
 - 26: gb_est26:*
 - 27: gb_est27:*
 - 28: gb_est28:*
 - 29: gb_est29:*
 - 30: gb_est30:*
 - 31: gb_est31:*
 - 32: gb_est32:*
 - 33: gb_est33:*
 - 34: gb_est34:*
 - 35: gb_est35:*
 - 36: gb_est36:*
 - 37: gb_est37:*
 - 38: gb_est38:*
 - 39: gb_est39:*
 - 40: gb_est40:*
 - 41: gb_est41:*
 - 42: gb_est42:*
 - 43: gb_est43:*
 - 44: gb_est44:*
 - 45: gb_est45:*
 - 46: gb_est46:*
 - 47: gb_est47:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
259: gb_est190:*
260: gb_est191:*
261: gb_est192:*
262: gb_est193:*
263: gb_est194:*
264: gb_est195:*
265: gb_est196:*
266: gb_est197:*
267: gb_est198:*
268: gb_est199:*
269: gb_est200:*
270: gb_est201:*
271: gb_est202:*
272: gb_est203:*
273: gb_est204:*
274: gb_est205:*
275: gb_est206:*
276: gb_est207:*
277: gb_est208:*
278: gb_est209:*
279: gb_est210:*
280: gb_est211:*
281: gb_est212:*
282: gb_est213:*
283: gb_est214:*
284: gb_est215:*
285: gb_est216:*
286: gb_est217:*
287: gb_est218:*
288: gb_est219:*
289: gb_est220:*
290: gb_est221:*
291: gb_est222:*
292: gb_est223:*
293: gb_est224:*
294: gb_est225:*
295: gb_est226:*
296: gb_est227:*
297: gb_est228:*
298: gb_est229:*
299: gb_est230:*
300: gb_est231:*
301: gb_est232:*
302: gb_est233:*
303: gb_est234:*
304: gb_est235:*
305: gb_est236:*
306: gb_est237:*
307: gb_est238:*
308: gb_est239:*
309: gb_est240:*
310: gb_est241:*
311: gb_est242:*
312: gb_est243:*
313: gb_est244:*
314: gb_est245:*
315: gb_est246:*
316: gb_est247:*
317: gb_est248:*
318: gb_est249:*
319: gb_est250:*
320: gb_est251:*
321: gb_est252:*
322: gb_est253:*
323: gb_est254:*
324: gb_est255:*
325: gb_est256:*
326: gb_est257:*
327: gb_est258:*
328: gb_est259:*
329: gb_est260:*
330: gb_est261:*
331: gb_est262:*
332: gb_est263:*
333: gb_est264:*
334: gb_est265:*
335: gb_est266:*
336: gb_est267:*
337: gb_est268:*
338: gb_est269:*
339: gb_est270:*
340: gb_est271:*
341: gb_est272:*
342: gb_est273:*
343: gb_est274:*
344: gb_est275:*
345: gb_est276:*
346: gb_est277:*
347: gb_est278:*
348: gb_est279:*
349: gb_est280:*
350: gb_est281:*
351: gb_est282:*
352: gb_est283:*
353: gb_est284:*
354: gb_est285:*
355: gb_est286:*
356: gb_est287:*
357: gb_est288:*
358: gb_est289:*
359: gb_est290:*
360: gb_est291:*
361: gb_est292:*
362: gb_est293:*
363: gb_est294:*
364: gb_est295:*
365: gb_est296:*
366: gb_est297:*
367: gb_est298:*
368: gb_est299:*
369: gb_est300:*
370: gb_est301:*
371: gb_est302:*
372: gb_est303:*
373: gb_est304:*
374: gb_est305:*
375: gb_est306:*
376: gb_est307:*
377: gb_est308:*
378: gb_est309:*
379: gb_est310:*
380: gb_est311:*
381: gb_est312:*
382: gb_est313:*
383: gb_est314:*
384: gb_est315:*
385: gb_est316:*
386: gb_est317:*
387: gb_est318:*
388: gb_est319:*
389: gb_est320:*
390: gb_est321:*
391: gb_est322:*
392: gb_est323:*
393: gb_est324:*
394: gb_est325:*
395: gb_est326:*
396: gb_est327:*
397: gb_est328:*
398: gb_est329:*
399: gb_est330:*
400: gb_est331:*
401: gb_est332:*
402: gb_est333:*
403: gb_est334:*
404: gb_est335:*
405: gb_est336:*
406: gb_est337:*
407: gb_est338:*
408: gb_est339:*
409: gb_est340:*
410: gb_est341:*
411: gb_est342:*
412: gb_est343:*
413: gb_est344:*
414: gb_est345:*
415: gb_est346:*
416: gb_est347:*
417: gb_est348:*
418: gb_est349:*
419: gb_est350:*
420: gb_est351:*
421: gb_est352:*
422: gb_est353:*
423: gb_est354:*
424: gb_est355:*
425: gb_est356:*
426: gb_est357:*
427: gb_est358:*
428: gb_est359:*
429: gb_est360:*
430: gb_est361:*
431: gb_est362:*
432: gb_est363:*
433: gb_est364:*
434: gb_est365:*
435: gb_est366:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634

Email: rdean@clemson.edu
Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 408.

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital status</i>	3. <i>Marital status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political affiliation</i>	8. <i>Political affiliation</i>
9. <i>Health status</i>	9. <i>Health status</i>
10. <i>Travel history</i>	10. <i>Travel history</i>
11. <i>Family size</i>	11. <i>Family size</i>
12. <i>Home ownership</i>	12. <i>Home ownership</i>
13. <i>Employment status</i>	13. <i>Employment status</i>
14. <i>Language spoken at home</i>	14. <i>Language spoken at home</i>
15. <i>Number of children</i>	15. <i>Number of children</i>
16. <i>Years in current residence</i>	16. <i>Years in current residence</i>
17. <i>Number of previous residences</i>	17. <i>Number of previous residences</i>
18. <i>Years since last move</i>	18. <i>Years since last move</i>
19. <i>Number of pets</i>	19. <i>Number of pets</i>
20. <i>Years since last pet acquisition</i>	20. <i>Years since last pet acquisition</i>
21. <i>Number of vehicles</i>	21. <i>Number of vehicles</i>
22. <i>Years since last vehicle purchase</i>	22. <i>Years since last vehicle purchase</i>
23. <i>Number of children under 18</i>	23. <i>Number of children under 18</i>
24. <i>Years since last child birth</i>	24. <i>Years since last child birth</i>
25. <i>Number of siblings</i>	25. <i>Number of siblings</i>
26. <i>Years since last sibling birth</i>	26. <i>Years since last sibling birth</i>
27. <i>Number of grandchildren</i>	27. <i>Number of grandchildren</i>
28. <i>Years since last grandchild birth</i>	28. <i>Years since last grandchild birth</i>
29. <i>Number of nieces/nephews</i>	29. <i>Number of nieces/nephews</i>
30. <i>Years since last niece/nephew birth</i>	30. <i>Years since last niece/nephew birth</i>
31. <i>Number of great-grandchildren</i>	31. <i>Number of great-grandchildren</i>
32. <i>Years since last great-grandchild birth</i>	32. <i>Years since last great-grandchild birth</i>
33. <i>Number of great-nieces/nephews</i>	33. <i>Number of great-nieces/nephews</i>
34. <i>Years since last great-niece/nephew birth</i>	34. <i>Years since last great-niece/nephew birth</i>
35. <i>Number of great-great-grandchildren</i>	35. <i>Number of great-great-grandchildren</i>
36. <i>Years since last great-great-grandchild birth</i>	36. <i>Years since last great-great-grandchild birth</i>
37. <i>Number of great-great-nieces/nephews</i>	37. <i>Number of great-great-nieces/nephews</i>
38. <i>Years since last great-great-niece/nephew birth</i>	38. <i>Years since last great-great-niece/nephew birth</i>
39. <i>Number of great-great-great-grandchildren</i>	39. <i>Number of great-great-great-grandchildren</i>
40. <i>Years since last great-great-great-grandchild birth</i>	40. <i>Years since last great-great-great-grandchild birth</i>
41. <i>Number of great-great-great-nieces/nephews</i>	41. <i>Number of great-great-great-nieces/nephews</i>
42. <i>Years since last great-great-great-niece/nephew birth</i>	42. <i>Years since last great-great-great-niece/nephew birth</i>
43. <i>Number of great-great-great-great-grandchildren</i>	43. <i>Number of great-great-great-great-grandchildren</i>
44. <i>Years since last great-great-great-great-grandchild birth</i>	44. <i>Years since last great-great-great-great-grandchild birth</i>
45. <i>Number of great-great-great-great-nieces/nephews</i>	45. <i>Number of great-great-great-great-nieces/nephews</i>
46. <i>Years since last great-great-great-great-niece/nephew birth</i>	46. <i>Years since last great-great-great-great-niece/nephew birth</i>
47. <i>Number of great-great-great-great-great-grandchildren</i>	47. <i>Number of great-great-great-great-great-grandchildren</i>
48. <i>Years since last great-great-great-great-great-grandchild birth</i>	48. <i>Years since last great-great-great-great-great-grandchild birth</i>
49. <i>Number of great-great-great-great-great-nieces/nephews</i>	49. <i>Number of great-great-great-great-great-nieces/nephews</i>
50. <i>Years since last great-great-great-great-great-niece/nephew birth</i>	50. <i>Years since last great-great-great-great-great-niece/nephew birth</i>
51. <i>Number of great-great-great-great-great-great-grandchildren</i>	51. <i>Number of great-great-great-great-great-great-grandchildren</i>
52. <i>Years since last great-great-great-great-great-great-grandchild birth</i>	52. <i>Years since last great-great-great-great-great-great-grandchild birth</i>
53. <i>Number of great-great-great-great-great-great-nieces/nephews</i>	53. <i>Number of great-great-great-great-great-great-nieces/nephews</i>
54. <i>Years since last great-great-great-great-great-great-niece/nephew birth</i>	54. <i>Years since last great-great-great-great-great-great-niece/nephew birth</i>
55. <i>Number of great-great-great-great-great-great-great-grandchildren</i>	55. <i>Number of great-great-great-great-great-great-great-grandchildren</i>
56. <i>Years since last great-great-great-great-great-great-great-grandchild birth</i>	56. <i>Years since last great-great-great-great-great-great-great-grandchild birth</i>
57. <i>Number of great-great-great-great-great-great-great-nieces/nephews</i>	57. <i>Number of great-great-great-great-great-great-great-nieces/nephews</i>
58. <i>Years since last great-great-great-great-great-great-great-niece/nephew birth</i>	58. <i>Years since last great-great-great-great-great-great-great-niece/nephew birth</i>
59. <i>Number of great-great-great-great-great-great-great-great-grandchildren</i>	59. <i>Number of great-great-great-great-great-great-great-great-grandchildren</i>
60. <i>Years since last great-great-great-great-great-great-great-great-grandchild birth</i>	60. <i>Years since last great-great-great-great-great-great-great-great-grandchild birth</i>
61. <i>Number of great-great-great-great-great-great-great-great-nieces/nephews</i>	61. <i>Number of great-great-great-great-great-great-great-great-nieces/nephews</i>
62. <i>Years since last great-great-great-great-great-great-great-great-niece/nephew birth</i>	62. <i>Years since last great-great-great-great-great-great-great-great-niece/nephew birth</i>
63. <i>Number of great-great-great-great-great-great-great-great-great-grandchildren</i>	63. <i>Number of great-great-great-great-great-great-great-great-great-grandchildren</i>
64. <i>Years since last great-great-great-great-great-great-great-great-great-grandchild birth</i>	64. <i>Years since last great-great-great-great-great-great-great-great-great-grandchild birth</i>
65. <i>Number of great-great-great-great-great-great-great-great-great-nieces/nephews</i>	65. <i>Number of great-great-great-great-great-great-great-great-great-nieces/nephews</i>
66. <i>Years since last great-great-great-great-great-great-great-great-great-niece/nephew birth</i>	66. <i>Years since last great-great-great-great-great-great-great-great-great-niece/nephew birth</i>
67. <i>Number of great-great-great-great-great-great-great-great-great-great-grandchildren</i>	67. <i>Number of great-great-great-great-great-great-great-great-great-great-grandchildren</i>
68. <i>Years since last great-great-great-great-great-great-great-great-great-great-grandchild birth</i>	68. <i>Years since last great-great-great-great-great-great-great-great-great-great-grandchild birth</i>
69. <i>Number of great-great-great-great-great-great-great-great-great-great-nieces/nephews</i>	69. <i>Number of great-great-great-great-great-great-great-great-great-great-nieces/nephews</i>
70. <i>Years since last great-great-great-great-great-great-great-great-great-great-niece/nephew birth</i>	70. <i>Years since last great-great-great-great-great-great-great-great-great-great-niece/nephew birth</i>
71. <i>Number of great-great-great-great-great-great-great-great-great-great-great-grandchildren</i>	71. <i>Number of great-great-great-great-great-great-great-great-great-great-great-grandchildren</i>
72. <i>Years since last great</i>	

```

1. .802
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgx80019C01r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/notes="Vector: pBACWICH; Site:1; HindIII;
Rice blast is one of the most devastating
of rice world wide. It is a filamentous
haploid genome (n=7) of approximately
blast is an important model fungal pathogen
numerous aspects of the fungal-host interactions
containing 9216 clones with an average insert
kbp was constructed. This library represents
than 25X genome coverage. High density

```

BASE COUNT	166 a	218 c	209 g	208 t	1 others
ORIGIN	are available upon request."				

Query Match	10.7%;	Score 150.6;	DB 227;	Length 802;
Best Local Similarity	56.7%;	Pred. No. 1.1e-33;		
Matches 301; Conservative	0;	Mismatches 224;	Indels 6;	Gap 1;

833	QY	tgctccattctgtgttggcttcactccagcaagaaatgatccaatacacactacttcgaaa	892
531	Db	TGTGCGAGTTCTGCGACGCTGTTTACGCAACGAGACTGGGAGGCATATGACTATCTCCAGA	472
893	QY	gcttgggttaagtactacggtttacggtgctggtgttaacccattgggtccagctcaagtggttg	952
471	Db	CACTGGGGAAGTGGATATGGTTACGCAATGGCAACCCCTGGCTCCACGCAAGGGGTGG	412
953	QY	gtttcgctaacgaattgattgcttagatgactcactctccaggttcaagacacacacttcta	1012
411	Db	GCTTCGTCAACGAGCTCATCGGAGGCTGCTCCAAAAGCCGTTGAGAGACCACAAATA	352
1013	QY	ctaacacacacttggactctaacccagctactttcccattgaaacgtactttgtcacgtg	1072
351	Db	CCAACTCGACGCTCGACAGCGACCCATCGAGCTTCCCACTAGACAAAAGCTGTACGCCG	292
1073	QY	acttctctcaagacacactatgatatactatttcttcgttttgggtttgtacaacggtta	1132
291	Db	ACTTTAGCCATGATAACGATATGCTTGGGCATCTACGCGCGCTGGGATTTACAACGCCA	232
1133	QY	ccaagccattgctactactcttctgtaatactattgaagaaactga-----cggttact	1186
231	Db	CGGCCCCGATTCGGTCCCAAAAAGAGAGAGAGCGCGCAGGAGCTCAGCGGGTCTCT	172
1187	QY	ctgcttcttgacactgttcattcgtctctagagcttactggttgaaatgatcaatgctcaag	1246
171	Db	CGTCCAGCTGGGGGATACCGTTCCGACGAGGATGTTTGTGTGAAAAAATGACTTGGCAG	112
1247	QY	ctgaaaaggaaccattgggttagagtttggtaacgacagagttgttccattgcacggtt	1306
111	Db	GGCAGACGAGGACTTGTGGAATCTTGTTCAACGACAGGGTTCACGCCGCTCGAACAAT	52
1307	QY	tgctgttgacaadtgtgggttagatgtgaagagacgacttcgttgaaggtt	1357

Description

Description	
AQ3244539	mgxb00019C
AQ3254540	mgxb00014M
AQ2554549	mgxb00017A
AQ2754007	AJ274007
AQ1630004	mgxb00021D
AQ1620040	mgxb00010F
AQ1611556	mgxb00008K
AQ3611495	mgxb00004K
AQ3611474	mgxb00004I
AQ202816	RPC111-48
AV1294327	AV129427
AV0641132	AV064132
AV2936463	AV293643
AV002759	AV002759
AV0771212	AV077132
AV114552	AV114552
AV081956	AV081956
AV060181	AV060181
AV053884	AV053884
AV118366	AV118366
AV047696	AV047696
AV3334544	AV333454
AV049962	AV049962
AV212676	AV212676
AV054552	AV054552
AV0140737	HS_3124_A
AV1287459	Tetraodon
AV161297	AV161297
AV054718	AV054718
AV056003	AV056003
AV091633	AV091633
AV049564	AV049564
AV079904	AV079904
AV078019	AV078019
AV100697	AV100697
AV120593	AV120593
AV055109	AV055109
AV053894	AV053894
AV151459	AV151459
AV058485	AV058485
AV050799	AV050799
AV067677	AV067677
AV054136	AV054136
AV118481	AV118481
AV164971	AV164971
AV054474	AV054474

RESULT 1
AQ324539/C

Accession	Size (bp)	Library	Source	Date
AQ324539	802	DNA	GSS	08-JAN-1999
mgxb0019C01r		CUGI Rice Blast	BAC Library Magnaporthe grisea genomic clone mgxb0019C01r, DNA sequence.	

AQ324539 .
AQ324539.1 GI:4116391

GSS.
Magnaporthe oryzae

Magnaporthe grisea

Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Yu, Y., Zhu, H., Boyd,

Phillips, K., Sasiniowski, M., Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the

Genome sequencing framework to sequence the Magnaporthe grisea genome

QY 442 aagcttggtagaaa---gattgttccattcattagagcttctggttgcagagatt 498
 Db 460 CAGAGTTAGCGGAGGATTCACCAATCCCTTTGTCAGCTCGGCTCTGAAGAGTC 519
 QY 499 attgcttctgtgaaagattcattgaaggtttcccaattctgctaag 543
 Db 520 GTCATGTGAGCGCAAGATTTTGTTCATGGCTCTCAAGAGCCAG 564

RESULT 4
 LOCUS AQ163004 753 bp DNA GSS 09-SEP-1998
 DEFINITION mgxb0021D19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0021D19r, DNA sequence.
 ACCESSION AQ163004
 VERSION AQ163004.1 GI:3559405
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 753)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 448.
 Location/Qualifiers
 1. .753

FEATURES
source

/organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0021D19r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site:1: HindIII; Site:2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."
 208 a 180 c 156 g 206 t 3 others

BASE COUNT
ORIGIN

Query Match 5.8%; Score 80.8; DB 225; Length 753;
 Best Local Similarity 48.8%; Pred. No. 7.le-13;
 Matches 274; Conservative 0; Mismatches 280; Indels .7; Gaps 2;

QY 90 ttgtgacactgttgacgttggttaccatgtttcccgaaatcttctcactgtgggtac 149
 Db 189 TTTTATGTCGCGCAGCCCGGCTTTCATTTTAAGACGCAATCACTATATATGGGGCCA 248
 QY 150 ctactctccactctctcttttggcagacgaatctgctatttctccagacgttccagac-- 207
 Db 249 GTATGACCACTATTTTGTGTCACCGTGCAGCCATCGGATTTATGATTCCTCTGCTACTT 308

QY 208 ----gactgtagagttacttctggttcaagtttctagacacggtgctagatacccaac 263
 Db 309 GCACAGGTTGATGATTTACGCTTTGCTCCCAAGTCTTTCAGAGGCACGGGGCTCGATACCCAAC 368
 QY 264 ttctctcgctctaaagcttactctgtcttggattgaagctattcaaaagacgctactgc 323
 Db 369 CGCTCAAAACCGGGCGGAAATTTCTCGAACACGATTCACCGACTCCAACCTCACGTACCGG 428
 QY 324 tttaagggtaagtagcgtttcttgaagacttacaactacacttgggtgctgcagactt 383
 Db 429 TAGTGCCTTATTGAATACTACATTAAAAAATACAAATCTCGGAGTCGAGGAATN 488
 QY 384 gactcattcgttgaacacaaatgtttaactctgtgtattaaagttctacaagaagataca 443
 Db 489 GAATGATTTGGCGCGCTCGGCAACAGAGAACTCCCGGTATTNATTTTACCAAGATACCA 548
 QY 444 ggccttggctagaagaattgttccattcatttagagcttctggttctgcagagttattgc 503
 Db 549 AAATCTGGCCAGAGAAAGCAACCATTTATTGCGTATGATGACAAANAACCGGTGTTGA 608
 QY 504 ttctgtcaaaagttcattgaaggtttccaatctgctaaagttgctgacccaggttctca 563
 Db 609 TAGCGCAGAACTCTCGGCTAGGGGTTT-CATCGAGCTTCTTATACAGATAAAGCGCGGC 667
 QY 564 accacacaaagcttctccagttattacagtgatcattccagaaggtccggttacaacaa 623
 Db 668 TAGACGAGAAACCTTTCCCTACAAAGCTGTGGCGTCTCTCACCAGCATGGGTTTATAA 727
 QY 624 cactttggaccacggttacttg 644
 Db 728 TACCTTTGACAAATAAGCTTG 748

RESULT 5

LOCUS AQ162040/c 699 bp DNA GSS 09-SEP-1998
 DEFINITION mgxb0010F06r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0010F06r, DNA sequence.
 ACCESSION AQ162040
 VERSION AQ162040.1 GI:3558441
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 699)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 285.
 Location/Qualifiers
 1. .699

FEATURES
source

/organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0010F06r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site:1: HindIII; Site:2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."

BASE COUNT	196 a	219 c	188 g	207 t
ORIGIN				

```

Query Match      5.3%; Score 74; DB 225; Length 810;
Best Local Similarity 58.0%; Pred. No. 7.8e-11;
Matches 131; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```

1179 cggttactctgcttcttggaactgttccattcgtctgtagcttagcttgaatgatgca 1238
QY
650 CGGGTCTCGTCAGCTGGCGGGTACCGTTCGCAGCGAGGATGTTTGTGAAAAATGAC 591
pb

1239 atgtoaagctgaaaaaggaaccatttggttagagtttggctaacgcacagagtgttccatt 1298
QY
590 ttgcgcaagggcacaacgacaggagctgttgagaatccttgctaacgcacagaggtgaccccgct 531
pb

[illegible]

QY 1359 gctttcgc tagatc tgggtaactgggc tgaagt ttcgcttaa 1404

RESULT 7

LOCUS	AQ361495	669 bp	DNA	GSS	03-FEB-1999
DEFINITION	mgxb0004kl0r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0004kl0r, DNA sequence.				

VERSION AQ361495.1 GI:4211334
KEYWORDS GSS.
SOURCE Magnaporthe oryzae.

ORIGIN: Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe; 1 (bases 1 to 669)

<p>AUTHORS</p> <p>Phillips, R. K., Sasnowski, M., Wing, R. A. and Dean, R. A.</p>	<p>TITLE</p> <p>A BAC End Sequencing Framework to Sequence the Magnaporthe grisea</p>
--	--

JOURNAL COMMENT
Unpublished (1998)
Contact: Dean RA
Clemson University Genomics Institute
Clemson, SC 29634-0901

100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293

Seq primer: GGAAACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 465.

```
1. .669
/organism="Magnaporthe grisea"
/strain="70-15"
source
```

```

/clone="mgxb0004kl0r"
/collection="CUGI Rice Blast BAC Library"
/clone_lib="CUGI Rice Blast BAC Library"
/clone_type="Protocl asts"

```

/lab_host= E. coli/Burk
 /note=vector PBACWICH; Site_1: HindIII; Site_2: HindIII
 most devastating fungal diseases
 Rice blast is one of the
 of rice. It
 a filamentous ascomycete with

RESULT 12
AV293643

us-09-488-265-28.rst

Sat Oct 27 15:25:46 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 17:46:43 ; Search time 3927.32 Seconds
(without alignments)
5529.654 Million cell updates/sec

Title: US-09-488-265-28
Perfect score: 1404
Sequence: 1 atggcgctgttcgtctgtct.....ggcgctgaatgtttcgttaa 1404

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_to.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vl1.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rol.*
95: gb_rol.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1360.8	96.9	1426	9	AX021809	AX021809 Sequence
2	798.8	56.9	1350	56	AF295325	AF295325 Synthetic
3	635	45.2	1404	9	A19452	A19452 phytase cdn
4	635	45.2	1404	10	I13430	I13430 Sequence 33
5	635	45.2	1404	10	I33881	I33881 Sequence 19
6	631	44.9	2000	13	ANPHYAG	Z16414 A.niger phy
7	631	44.9	2665	13	ASNPHYTASE	M94550 Aspergillus
8	631	44.9	6756	9	A19451	A19451 phytase gen

Sat Oct 27 15:25:44 2001

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1404)
AUTHORS Van Ooijen,A.J.J., Rietveld,K., Hoekema,A., Pen,J., Sijmons,P.C.
and Verwoerd,T.C.
TITLE Expression of phytylase in plants
JOURNAL Patent: US 5593963-A 19 14-JAN-1997;
FEATURES Location/Qualifiers
source 1..1404
/organism="unknown"
BASE COUNT 293 a 436 c 344 g 331 t
ORIGIN

Query Match 45.2% Score 635; DB 10; Length 1404;
Best Local Similarity 65.8%; Pred No. 1.3e-161;
Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;

QY 1 atggcggtgttcgtgctactgtccattgccacccctgttgcgttccacacacgcgtacc 60
DB 1 ATGGCGGTCTGCTGTTACTTCTTGTATCTCCTGCTGCGAGTCACTCCGGACTG 60
QY 61 gcttgggtctcgtggttaattcactcttctgtacactgttgacggtgttacaatgt 120
DB 61 GCAGTCCCGCGCTCGAGAAATCAATCCAGTTGCGATACGGTTCGATCAGGGGTATCAATGC 120
QY 121 tccccagaaatttcactgttgggttacctactctccactctctcttcttggcgagcgaa 180
DB 121 TTCTCCGAGACTTCGACTCTTGGGTCAATACGCACCGTTCCTCTCTCTGCGCAACGAA 180
QY 181 tctgtatttctccagacgttccagacgaactgttagagtaactctcgttaagtttgtct 240
DB 181 TCGGTTCATCTCCCTGAGTGTGCGCGCGGATCGAGTCACTTCCTCGTCAAGTCTCTCC 240
QY 241 agacaggtgttagataccacacttctcgtctcgaagcttactctgtctgttgatgaa 300
DB 241 CGTCATGGAGCGCGGTATCCGACCGACTCCCAAGGCAAGAAATACCTCGCTCATTTAG 300
QY 301 gctattcaaaagacgtactgtcttcaagggttaagtaagtaacgtcttcttgaagacttaaac 360
DB 301 GAGATCCACGACACGCGACCCACTTTGAGGAAATATGCTTCTCTGAAGACATACAAAC 360
QY 361 tacaatttgggtgtgacagactgtactccactcgttgtaaaacaaatgttaactcttgt 420
DB 361 TACAGCTTGGGTGCAATGACCTGCTGCGAGAACAGAGCTAGTCAATCCCGC 420
QY 421 attaaattctacagaatatacagaagcttgggtgtagaagaattgttccattcattagact 480
DB 421 ATCAAGTTTACCAGCGGTACGAAATCGCTCAAGGAACATCGTTCCATTATCCGATCC 480
QY 481 tctggttctgacaggtattgtctctgctgaaagttcaatgaaagtttccaattctgt 540
DB 481 TCTGGCTCAGCGCGTGTATCGCTCCGCAAGAAATTCATCGAGGGCTTCCAGAGCAC 540
QY 541 aagttggtgacccaagttctcaacacacacacacacacacacacacacacacacacacac 600
DB 541 AAGCTGAAGGATCTCTGCTCCCGACCGCGCCAAATGCTCGCCCAAGATCGAGTGTCTAT 600
QY 601 ccagaaggtatccggtttac 660
DB 601 TCCGAGGCCAGCTCATCCAAACACTCTCGACCCAGGACCCCTGCTCTTCGAAGAC 660
QY 661 tctgaattaggtgacagcttgaagtaacttaactcactgttctgttgcctccagctttaga 720
DB 661 AGCGAATTGGCCGATACCGTCGAAGCAATTTACCGCCGACGCTGCTCCCTCCATTCGT 720
QY 721 gctagattggaagctgacttgcaggtgttaactttgactgaagaagacactgttactgt 780
DB 721 CAACGCTGGAGAACGACCTGTCCGGTGTGACTCTCACAGACACAGAACTGACCTACCTC 780
QY 781 atggacatgtgtccattcgacactgtcgttagaacttctgaagcgtactgaattgtctcca 840

DB 781 ATGGACATGTCTCTCTCGACACCATCTCCACAGCACCGTCGACACCAAGCTGTCCCCC 840
QY 841 tctgtgttcttctcactcagcagaataggtatcaatacactactgttgcgaagcttgggt 900
DB 841 TTCTGTGACCTGTTTCAACCATGAGCAATGATGAATACAGTACTACCTCCAGTCTTGA 900
QY 901 aagtaactacgtttagcgtggtgtaacccattgggtccagctcaaggtgttgggttgcgt 960
DB 901 AAGTATTACGGCCATGGTGCAGGTAAACCCGCTGCGCCGACCCAGGGGCTCGGCTAGCT 960
QY 961 aacgaattgattgtagattgactcactctccagttcaagaccacacacttctactaacac 1020
DB 961 AACGAGCTCATCGCCGCTGTACCCACTCGCTGTCCAGCATGACACACAGTTCACACAC 1020
QY 1021 actttgactctaacccagctacttctccattgaacgtacttctgtcgtgactctctct 1080
DB 1021 ACTTTGGACTCGAGCCCGGCTACCTTTCCGCTCAACTCTACTCTCTACCGCGACTTTTCG 1080
QY 1081 cagcaacacactatgatattcttcttcttcttgggtttgtgacaacgggtaccagcca 1140
DB 1081 CATGACACCGCATCATCTCCATCTCTTTGCTTTAGGTCTGTACACGGCACTAAGCCG 1140
QY 1141 tigtctactactctotgttgaatctattgaagaactgaacgttactctgtcttcttggact 1200
DB 1141 CTATCTACCCACGACCGTGGAGAAATATACCCAGACAGATGGATTCTCTGCTCTTGGAGC 1200
QY 1201 gttccattcgtcgttagagcttactgttgaatgatgcaatgtcaagctgaagaagcca 1260
DB 1201 GTTCCGTTTCTGCGGTTTGTACGTCGAGATGATGCACTGTGTCAGCGGAGCAGGCGG 1260
QY 1261 ttggttagagtttgggttaacgacagagttgttccattgacaggttctgtctgtgacaag 1320
DB 1261 CTGCTCGGCTCTGCTGTTAATGATCGCTGTGTCGCTGTCATGGGTGTCGCTTGTATGT 1320
QY 1321 ttgggttagatgttaagagacagactcgttgaagtttcttctcgttagactctgttgt 1380
DB 1321 TTGGGAGATGTACCCGGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGT 1380
QY 1381 aactgggctgaattgttctgctta 1403
DB 1381 GATTGGCGGAGTGTGTTGCTTA 1403

RESULT 6
ANPHYAG ANPHYAG 2000 bp DNA PLN 06-OCT-1992
LOCUS A.niger phyA gene.
DEFINITION Z16414
ACCESSION Z16414.1 GI:2392
VERSION Z16414.1 GI:2392
KEYWORDS phyA gene.
SOURCE Aspergillus niger.
ORGANISM Aspergillus niger
Aspergillus niger
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE 1 (bases 1 to 2000)
AUTHORS van Hartingsveldt,W., Van Zeijl,C.M.J., Harteveld,M.G., Gouka,R.J.,
Suykerbuyk,M.E.G., Luiten,R.G.M., Van Paridon,P.A., Sellen,G.C.M.,
Veenstra,A.E., Van Gorcom,R.F.M. and Van Den Hondel,C.A.M.J.
TITLE Cloning, molecular characterization and overexpression of the
phytase gene (phyA) of Aspergillus niger
JOURNAL Gene (1992) in press
REFERENCE 2 (bases 1 to 2000)
AUTHORS van Hartingsveldt,W.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-1992) Van Hartingsveldt W., TNO Medical
Biological Laboratory, Lange Kleiweg 139, Rijswijk, the Netherlands
FEATURES
source 1..2000
/organism="Aspergillus niger"
/strain="Van Rieghem strain"
/db_xref="taxon:5061"
TATA_signal 131..134

exon	166..253 /gene="phyA" /number=1	744	CAGCGGTACGAATCGTTCACAAGGAACATCGTTCCATTCCATCCGATCCTCTGCTCCAGC	803
gene	166..1871 /gene="phyA"	493	agagttattgtctctgctgtaaaagtttcattgaaggtttccaaatctgctgaagtgtgctgac	552
prim_transcript	166..1871 /partial	804	CGCGTGATCCCTCCGCCAACAAATTCATCGAGGGCTTCCAGAGCACCAGCTGAAGGAT	863
sig_peptide	166..1871 /gene="phyA"	553	ccaggttctcaaccacacaaagcttctccagttattaaagtgatcatctccagaagatcc	612
CDS	166..1871 /function="primary transcript of phyA gene" /gene="phyA"	864	CCTCGTCCCGAGCCGCGCCAAATCGTCGCCCAAGATCGAGTGGTCTATTTCCGAGGCCAGC	923
	166..1871 /gene="phyA"	613	ggttacaacaacactttggacaacggtactgttactgcttcttogaagaactctgaattaggt	672
	166..1871 /codon_start=1	924	TCATCCAAACACACTCTCGACACCGACCTGCACCTGTCTTTCGAGACACAGCAATTTGGCC	983
	166..1871 /db_xref="GI:2393"	673	gacgagttgaagctaactcactgcttctgttcgctccagctattagactagattggaa	732
	166..1871 /translation="MGVSALLPLYLISGVTSGLAVPASRNSSCDTVQGYCFSET SHLQYAPFFSLANESVTSPEVAGCRVTFQVLSRHGARYPTDSKGRKYGALIBEI QQNATTFALFKLYNYSLEADDLTFCEQELVNSGIKFORYESLRNIVPFTSR SGSRVIAAGKFEIGFQSTKLMDRAQPGQSPKIDVVISSEASSNNLTLDPGCTVF EDELADTVEANFATFVPSIRORLENDLSGVTLTDEVTYLMDCSPDTISFVPTD KLSPFCDLTHDWINVLYQSLKYYGHGANGPLGPTGVGYANELIARLTHSPVHD DTSSNTLDSPPATFPLNSTLYADFSHDNGLIISILFALGLYNGTLPDLSPTVENTITQ DGSSAWTVPPASRLYVEMQCAEQEPLRVLVNDRVVLHGCPVDLGRCTRDSFV RGLSPARSGDNWAECA"	984	GATACCGTCGAAGCCAAATTCACCGCCACGTTTCGTCCTCCATTCGTCAACGTCFGGAG	104
intron	166..1871 /db_xref="GI:2393"	733	gctgacttgcaggtgttacttctgactgacgagacgtttgttacttgatgacatgtgt	792
	166..1871 /number=1	1044	ACGACCTGTCCGGGTGTACTCTCAGACACAGAAAGTACACCTACCTCAGACATGTGC	110
	166..1871 /number=2	793	ccattcgacactgtcgtgtagaactctgacgctactgactgaattgttcattctgtgctttg	852
	166..1871 /gene="phyA"	1104	TCCTTCGACACCATCTCCACACGACCGTCGACACCAAGCTGTCCCTCTTCGTGACCTG	116
	166..1871 /gene="phyA"	853	ttaactcacgaaatggatccaatacactacttgcgaaagcttggtaagtactacggt	912
	166..1871 /number=1	1164	TTACCCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTGAAAGATTTACGGC	122
mat_peptide	166..1871 /gene="phyA"	913	taogtgcgtgtaacccattgggtccagctcaagtggtgtgttgcgtacaagaattgatt	972
stem_loop	166..1871 /gene="phyA"	1224	CATGTTGACAGGTAAACCCGCTCGGCCCGACCCACGAGCGTCGGTACGCTAACGAGTCA	128
BASE COUNT	463 a 603 c 456 g 478 t	973	gctagattgactactctccagtttcaagacacacttctacttaacacacacttttgactet	103
ORIGIN		1284	GCCGCTGACCCACTCGCTGTCCACGATGACACCAAGTTCACACCAACACTTTGGACTCG	134
Query Match	44.9%; Score 631; DB 13; Length 2000;	1033	aaccagctacttcccaattgaaagctactttgtacgctgacttctccagcaacaact	109
Best Local Similarity	65.9%; Pred. No. 1.7e-160;	1344	AGCCCGGTACGCTTCCGCTCAACTCTACTCTACGCGGACTTTTCGCATGACACACGGC	140
Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;		1093	atgatatctatttctcgtcttgggtttgtacaacggtgacacgaactgtctactact	115
		1404	ATCATCTCAATCTCTTTGCTTTTAGTCTGTACAACGCACTAAAGCCGCTATCTACCA	146
		1153	ctgttgatctattgaagaactgacggttactctgcttcttgactgttgcactgtccattcgt	121
		1464	ACCGTGGAGATATCACCCACACAGATGGATTCGCTCTGCTTGGACGGTTTCCCGTTGCT	152
		1213	gctagagcttactgtaaatgatgcaatgtcaagctgaaaggaaccatttggtttagagtt	127
		1524	TCGGCTTTGTACGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	158
		1273	ttggttaacgacagagttgttccattgcacggttgcctgtgttgcacagagttggtagatgt	133
		1584	TTGTTAATGATCCGTTGTCCTCCCTGCATGGGTGTCCGGTTGATGCTTTGGGAGATGT	164
		1333	aagagacgacttcgttgaaggtttgtcttctgcgtagactgtgtgtaactgaggtgaa	139
		1644	ACCCGGATAGCTTTGTCGAGGGGTTGAGCTTTTCTAGATCTGGGGGTGATTGGCGGAG	170
		1393	tggttcgctta 1403	
		1704	TGTTTGTCTTA 1714	
RESULT	7			
ASNPHYTASE	2665 bp	DNA	PLN	27-APR-1993
LOCUS	Aspergillus niger myo-inositol hexaphosphate phosphohydrolyase g			
DEFINITION	complete cds.			
ACCESSION	M94550			

M94550.1 GI:166520
 myo-inositol hexaphosphate phosphohydrolyase: phytase.
 Aspergillus niger (library: EMBL 3; NRRL 3135) DNA.
 ORGANISM
 Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 1 (bases 1 to 2665)
 Mullane, E.J., Gibson, D.M. and Ullah, A.H.
 Positive identification of a lambda gtlI clone containing a region of
 fungal phytase gene by immunoprobe and sequence verification
 Appl. Microbiol. Biotechnol. 35, 611-614 (1991)
 JOURNAL
 MEDLINE
 REFERENCE
 2 (bases 1 to 2665)
 Mullane, E.J.
 Sequence of the Aspergillus niger (ficusum) phytase gene
 Unpublished (1992)
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..2665
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 /tissue_lib="EMBL 3; NRRL 3135"
 <683..726
 join(683..726,829..2188)
 /note="phytase gene"
 /codon_start=1
 /protein_id="AAA32705.1"
 /db_xref="GI:166521"
 /translation="MGVSALLPLLYLGSVSLVAPASRNQSCDVTGQYCFSET
 SHLWGOVAPFSLANESVISPAGCRVTFQVLSRHGARYPTDSKGRYSALIEEI
 QONATFDGKFLKTYNSIGADLLTFGEQELVNSGKIFQYRSLRNIPVFI
 SGSRVSLASGKKFLEGQSKLKDPRAGQSSPKIDVISEASSNNTLDPGCTVF
 EDELEADVEANETATFVPSQRKLENDLSGVLTDFEYVLMDCSEDFISSTVD
 KLSPFDLTHDEWINDYLOSRLKYGHGAGNPLGPTQGVGVANELLARLHSPVHD
 DTSNHTLDSSPATFLNMLYADFSHDNGIISLFDALYNGKPLUSTIVENITQT
 DGFSAWTVPEASRLYVEMMQCAEQEPLRVLVNDRVPLHGPCVDALGRCTRDSFV
 RGLFSRSGGWDACEFA"
 join(683..726,829..841)
 727..828
 829..2188
 842..2176
 /standard_name="phytase"
 /EC_number="3.1.3.8"
 /function="catalyzes the hydrolysis of phytic acid"
 /product="myo-inositol hexaphosphate phosphohydrolyase"
 BASE COUNT 633 a 808 c 574 g 650 t
 ORIGIN
 Query Match 44.9%; Score 631; DB 13; Length 2665;
 Best Local Similarity 65.9%; Pred. No. 1.7e-160;
 Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;
 QY 13 gtcgtgctactgtccattgccacttctgttcggttcacacatccggtaccgcttggtgct 72
 Db 797 GTGGGACTACTGCTGCTACTATCTGTCAGAGTCACCTCCGGACTGCGACGTCGCCGCGC 856
 QY 73 cgtggaattctcaactctgtgacactgttgacggtggttaccatgtttccacagaaat 132
 Db 857 TCAGAAATCAATCCAGTTCGATACGGTCATCAGGGGTATCAATGCTTCTCCGAGACT 916
 QY 133 tctcaactgtgggtaccactctccactctcttcttcttgggacacaaatctgctattct 192
 Db 917 TCGCATCTTTGGGTCAATACGACCGCTTCTTCTCTGCGCAACGAATCGGTCACTCC 976
 QY 193 ccagagcttccagacactgtagatgactcttctgacagtttctgacagcagctgct 252
 Db 977 CCTGAGTGCCTCCGCGATCGAGATCACTTTCCTCAGGTCTCTCCCGTCTATGGAGCG 1036
 QY 253 agatacccaacttcttctgctgcttaaggcttactctgttggattgaagcttcaaaag 312
 Db 1037 CGGTATCCGACCGACTCCAGGGCAAGAAATACTCGCTCTCTATTGAGGAGATCCAGCAG 1096
 QY 313 aacgctactgcttccagggtgaagtagcgtttcttgaagacttaacactattgggt 372

1097 AACGGACACACCTTTGACGGAAATATGCTTCCTGGAAGACATACAACTACAGCTGGGT 1156
 QY 373 gctgacgactgtactccatttcggtgaaacacaaatggttaactctggtattaaagttaac 432
 Db 1157 GCAGATGACCTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 1216
 QY 433 aagaatatacaaggcttggctgagaagattgttccattcattagagcttctggttctgac 492
 Db 1217 CAGCGGTACGAATCGCTCACAAGGAACATCGTTCATTCATCCGATCTCTGGCTCCAGC 1276
 QY 493 agagttattgcttctgtgaaagttcatttcgaaggtttccaatctgcttaagtgtgctac 552
 Db 1277 CGCGTATCGCTCCGGCAAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGGAT 1336
 QY 553 ccaggttctcaaccac 612
 Db 1337 CCTCGTCCGACGCCGCCAAATCGTCGCCCAAGATCGAGCTCGCTCCCTCCATTCGAGGCCAGC 1396
 QY 613 ggttaacaacaacacttggaccacacacacacacacacacacacacacacacacacacac 672
 Db 1397 TCATCCAAACAACACTCTCGACCCAGGCACCTGCACTGTCTTCGAAGACAGCAATGGCC 1456
 QY 673 gacgacttgaagacttaactcactgcttctgttcgctccagctattagagctagtgaa 732
 Db 1457 GATACCGTGAAGCCAAATTCACCGCCACGTTCTGTCGCCCTCCATTCGTCACGCTGGAG 1516
 QY 733 gctgacttgcaggttacttactgactgacgaagacgcttcttacttactgacgacatgct 792
 Db 1517 AACGACCTGTCGGGTGTGACTCTCACAGACACAGAGTCACTACCTCACTGACATGTCG 1576
 QY 793 ccatcgacactgctctgagaactctctgacgctactgaaatgtctctcattctgctgttg 852
 Db 1577 TCCTCGACACCATCTCCACACACACCGTCGACACCAAGCTGTCCCTCTCTGTGACCTG 1636
 QY 853 tcaactcacacgaatggaatccaaatagcacttctgaaagcttgggtgaagtactacgt 912
 Db 1637 TTCACCATGACGAATGGATCAACTACGACTACCTCCAGTCCCTTTGAAAAGTATACGGC 1696
 QY 913 tagcgtgctgtaacccattgggtccagctcctcaagggtgtgtgttctgctcaacgaattg 972
 Db 1697 CATGTCGACGTTAACCGCTCGCGCCGCCACCCAGGGCTCGGCTACGCTAACGAGCTCATC 1756
 QY 973 gctagattgactcactctccagttccaaagacacacacttctactaaacacacttggact 1032
 Db 1757 GCCGCTGTGACCCACTCGCTCTCCAGATGACACACAGTTCCAAACACACATTTGGGACTG 1816
 QY 1033 aaccagacttctccattgaaagcacttctgacgctgacttctcactcactcactcactcact 1092
 Db 1817 AGCCCGGCTACCTTTCGCTCAACTCTACTCTCTACGCGGACTTTTCGATGACAAACGGC 1876
 QY 1093 atgatatctatttcttctgcttgggtttgtacaaoggttaccagcattgcttactact 1152
 Db 1877 ATCATCTCCATCTCTTTGCTTGTAGTCTGTACAAACGACACTAAGCGCGTATCTACCA 1936
 QY 1153 tctgtgaatctattgaagaac 1212
 Db 1937 ACCGTGGAGAATATACCCGACAGATGGAATCTCGTCTCTGCTGGAGCGTTCGCTTGGCT 1996
 QY 1213 gctagagcttacgttgaaatgtagcagctgacgttctctctctctgactgctgactgct 1272
 Db 1997 TCGGTTGTACGTCGAGATGATGTCAGTGTGTCAGCGGAGGAGGAGCGCTGGTCCGTGTC 2056
 QY 1273 ttggttaacgacagagttgttccactgacaggttctgctgttgaacagttgggttagatgt 1332
 Db 2057 TTGTTAATGATCGCGTGTCTCCGCTGCATGGGTGTCCGGTGTGATGCTTTGGGGAGATGT 2116
 QY 1333 aagaagagcactctgcttgaaggtttgtcttctgctagatctggtgtaactgggctgaa 1392
 Db 2117 ACCCGGATAGCTTTGTCAGGGGGTGTGAGCTTGTCTAGATCTGGGGGTGATTGGCGGAG 2176
 QY 1393 tgtttcgttta 1403
 ||||| |||||


```

Db 2177 TGTGTTGCTTA 2187
RESULT 8
LOCUS A19451 6756 bp DNA PAT 10-JUN-1994
DEFINITION phytase gene locus of plasmids pAF 2-3, pAF 2-6, pAF 2-7.
ACCESSION A19451
VERSION A19451.1 GI:583193
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS van Gorcom, R.F.M., van Hartingsveldt, W., van Paridon, P.A.,
Veenstra, A.E., Luitzen, R.G.M. and Sellen, G.C.M.
TITLE Cloning and expression of microbial phytase
JOURNAL Patent: EP 0420358-A 40 03-APR-1991;
GIST-BROCADES N.V.
FEATURES
source i. 6756
/organism="synthetic construct"
/db_xref="taxon:32630"
join(210..254,357..1715)
/gene="phytase gene"
join(210..254,357..1715)
/gene="phytase gene"
join(210..254,357..1712)
/gene="phytase gene"
210..254
/gene="phytase gene"
/number=1
join(210..254,357..1715)
/gene="phytase gene"
/codon_start=1
/transl_table=1
/protein_id="CAA01463.1"
/db_xref="GI:583194"
/translation="MGSVAIVLLPLVLLSGVTSGLAIVASRNQSSCDTVDDGYQCFSET
SHWQVAPFFSLANESVPEVACGRVTEAQLSRHGARYPTDSKGGKYSALIEEI
QONATFQGLAFKATYNSILGADDLTPFGQELVNSGIFKYQRYESI,TRNVPFIRS
SGSRVIAAGKRFIEGFOSTKLPDRAPOGSSPKIDVYISEASSNNILDPGCTVF
EDSELADVEANFTATFVPSIKORLENDLSGVTLDTEVTYLMDCSFDTISTSTVDI
KLSPCDLFDHEWNYLDYLSKKYGHGAGNPLGTPQGVYANELIARLTHSPVHD
DPSNNHLDSPATFPLNSTLYADPFDHNGIISILFALGLYNGTKPLTIVENITQT
DGFSSAWIVPFASRLYVEMMQCAEQEPLRVLVNDRVPLHGPCVDALGRCTRDSFV
KGLSFARSGGDWAECFA"
255..356
/number=1
357..1715
/gene="phytase gene"
/number=2
BASE COUNT 1615 a 1712 c 1712 g 1714 t 3 others
ORIGIN
Query Match 44.9%; Score 631; DB 9; Length 6756;
Best Local Similarity 65.9%; Pred. No. 1.8e-160;
Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;
QY 13 gtcgtgactgtccattgccacctgttcggttccacacatccggtaccgccttggtct 72
Db 324 GTGGGACTACTGCTGACTATCTGTGCAGAGTCACCTCCGGACTGGCAGTCCCCGCC 383
QY 73 cgtgtaattctcactctgtgacactgttgacggtgttaccacatgtttcccgaaatt 132
Db 384 TCGAGAAATCAATTCAGTTGGCATACGCTGCATCAGGGGTATCAATGCTTCCGAGACT 443
QY 133 tctcactgtgggttacactctccataactctcttcttggcagacgaatctgtattct 192
Db 444 TCGCATCTTTGGGTCAATACGACCGCTCTCTCTCTGGCAACGAATCGGTATCTCC 503
QY 193 ccagacgttccagacgactgtagagttacttcttccagattgttctagacacggtgct 252

```

Ddb	1584	TTGGTTAATGAATCCGGTTGTCCCGCTGCATGGGTGCTCCGGTTGATGCTTTGGGAGATGT	1643
Qy	1333	aagagagacgacttcgttgaaggttctcttcgctagatctcgtgtaactcgggctgaa	1392
Ddb	1644	ACCGGGATAGCTTTGTGAGGGGTTTGAGCTTTGCTAGATCTGGGGTGTATTGGCGGAG	1703
Qy	1393	tgtttcgctta	1403
Ddb	1704	TGTTTTGCTTA	1714
RESULT	9		
LOCUS	II13429	6756 bp	DNA
DEFINITION	Sequence 31 from patent US 5436156.		PAT
ACCESSION	II13429		
VERSION	II13429.1	GI:910770	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 6756)		
AUTHORS	Van Gorcom,R.F.M., Van Hartingsveldt,W., Van Paridon,P.A., Veenstra,A.E., Luijten,R.G.M. and Sellen,G.C.M.		
TITLE	Cloning and expression of phytase from aspergillus		
JOURNAL	Patent: US 5436156-A 31 25-JUL-1995;		
FEATURES	Location/Qualifiers		
source	1..6756		
BASE COUNT	1615 a 1714 c 1712 g 1715 t		
ORIGIN	/organism="unknown"		
Query Match	44.9%;	Score 631;	DB 10; Length 6756;
Best Local Similarity	65.9%;	Pred. No. 1.8e-150;	
Matches 916;	Conservative 0;	Mismatches 475;	Indels 0; Gaps
Qy	13	gtcgtgtactgtccattgccacctgttcggttcacatcccgtagccgcttgggttcct	72
Ddb	324	GTGGGACTACTGATCGTGACTATCTGTGCAGAGTCACCTCGGACTGGCAGTCCCGGCC	383
Qy	73	cgtggtaatctcaactccttgtagaactgttgacggtgtgtaccaatgtttccagaatt	132
Ddb	384	TCGAGAAATCAATCCCAATTCGCATACGGTCGATCAGGGGTATCAATGCTTCTCCGAGACT	443
Qy	133	tctcactgtgggttaccttactctccatctctctctctctctctctctctctctctctct	192
Ddb	444	TCGCATCTTTGGGGTCAATACGCACCGTCTCTCTCTCTGGCAACGAATCGTCAATCC	503
Qy	193	ccagacggtttccagacgactgtagattccttgcctcgaagtttgcctagacacggtgct	252
Ddb	504	CCTGAGTGCCTCCCGGATGCAGAGTCACTTTTCGCTCAGGTCTCTCCCGTCATGGAGCG	563
Qy	253	agatacccaacttcttcgctcgtcgaagcttactcgtcttcttgattgaagctattcaaa	312
Ddb	564	CGGTATCCGACCGACTCCCAAGGCAAGAANAATCTCCGCTCTCATTCAGGAGATCCAGCAG	623
Qy	313	aacgctactgtcttcgaaggtgaagtacgcttcttcgaagacttaacaactacacttgggt	372
Ddb	624	AACGGACCACTTTGACGGAAAATATGCTTCTCCTGAAGACATACAACATACAGCTTGGGT	683
Qy	373	gctgaagacttgactccattcggtagaaaccaaaatggttaaactctggatttaagtctac	432
Ddb	684	GCAGATGACCTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGCATCAAGTTCATC	743
Qy	433	agaagatcacaaagccttggcctgaagaagattgttccattcattagagcttctgttctgac	492
Ddb	744	CAGCGTACCGAATCGCTCACAAAGACATCGTTCCATTTCATCCGATCTCTTGGCTCCAGC	803
Qy	493	agagtattctctcgtgaaagtttcattgaaggtttccaatctgctaagtctgctgac	552
Ddb	804	CGCGTATCGCTTCGGCAAGAAATTCATCGAGGGTCTCCAGGACCAACGCTGAAGGAT	863

10

us-09-488-265-28.rge

DEFINITION Sequence 7 from patent US 5780292.
 ACCESSION AR018076
 VERSION AR018076.1 GI:3973679
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2363)
 AUTHORS Nevalainen, H.K.M., Paloheimo, M.T., Miettinen-Oinonen, A.S.K.,
 Torkkeli, T.K., Cantrell, M., Piddington, C.S., Rambosek, J.A.,
 Turunen, M.K., and Fagerstrom, R.B.
 TITLE Production of phytate degrading enzymes in trichoderma
 JOURNAL Patent: US 5780292-A 7 14-JUL-1998;
 FEATURES Location/Qualifiers
 source 1..2363
 BASE COUNT 559 a 732 c 510 g 562 t
 ORIGIN
 Query Match 44.8%; Score 629.4; DB 9; Length 2363;
 Best Local Similarity 65.8%; Pred. No. 4.6e-160;
 Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
 QY 13 gctgctactgtccatgacacattgttgcggttccacatccggtacacgcttggtgct 72
 Db 518 GTGGGACTACTGATCGCTGACAATCTGTGCAGAGTCACCTCCGGACTGGCAGTCCCGGCC 577
 QY 73 cgtgtaattccactctgttgacactgttgacggtggtttaccatgtttcccgaaatt 132
 Db 578 TCGAAGATCAATCACTTCGATGCGTGCATCAAGGGTATCAATGCTTCTCCGAGACT 637
 QY 133 tctcactgtgggttacacttctccatctccatctcttcttcttcttcttcttcttcttct 192
 Db 638 TCGCATCTTTGGGTCAATACGCGCGGTCTCTCTCTGCGAAACGAATCGGCCATCTCC 697
 QY 193 ccagcgttccagacgactgtagactgttcttcttcttcttcttcttcttcttcttcttct 252
 Db 698 CTTGATGTCCCGCGGTGTCAGAGTCACTTTCGCTCAGGTCTTCCCGTCATGGAGCG 757
 QY 253 agatacccaactcttctgcttgaagcttactcttcttcttcttcttcttcttcttcttcttct 312
 Db 758 CGGTATCCCGACCGAGTCCCAAGGCAAGAAATCTCCGCTCTCATTTGAGGAGATCCAGCAG 817
 QY 313 aacgctactgtcttccaggttaagtaagcttcttcttcttcttcttcttcttcttcttcttct 372
 Db 818 AACGTGACCAACCTTTGATGAAATATGCTTCTTGAAGACATACAACTACAGTCTGGGT 877
 QY 373 gctgacgacttgaactcattcgtgtaaaacaaatggttaactctggttattaaagtctac 432
 Db 878 GCAGATGACTGACTCCCTTCGGAGAGCAGAGTAGTCAACTCCGGCATCAAGTTCTAC 937
 QY 433 agaagatacaaggttggctagaagattgttccattctattagagcttctggttctgac 492
 Db 938 CAGCGATACGAATCGTCAAGAGAACATCATTCGCTTCATCCGATCTCTGGCTCCAGC 997
 QY 493 agagtattgtctctgtaaaagttaaggttcttccatctgcttaagtgtgctgac 552
 Db 998 CGGTGATCCGCTCCGGGCAATTCATTTGAGGGTCTCCAGCAGCAGCTGAAGAT 1057
 QY 553 ccaggttctcaac 612
 Db 1058 CTTGTCGCCAGCGGGCCCAATCTGTCGCCCAAGATCGACGTGTGTCTATTTCCGGGCCAGC 1117
 QY 613 ggttcaacacacacttggaccacggtacttcttcttcttcttcttcttcttcttcttcttct 672
 Db 1118 TCATCCAAACAACACTTCGACCCAGGACCTGCACTGTCTTTGAACACACAGCAATTTGGCC 1177
 QY 673 gacgacttaagactaacttacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 732
 Db 1178 GATACCTGCAAGCAATTTCAACGCCACGCTTCGCCCTCCATTCGTAACGCTCTGGAG 1237
 QY 733 gctgacttgcaggtgttacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 792

Db 1238 AACGACCTGTCTGGCTGACTCTCACACACACAGAGTGACTACCTACCTCATGGACATGTC 1297
 QY 793 ccattcgcactgtctgtagaacttctgaagcttactgaattgtctcattctgtgcttg 852
 Db 1298 TCCTTTCGACACCATCTCCACACGACACCGCTGACCAAGCTGTCCCTCTCTGTGAGCTG 1357
 QY 853 ttcaactcagcagaatggtccaatcagactacitgaaagcttgggttaagtactacggt 912
 Db 1358 TTACCCCATGAGGAATGATCCATACGACTACCTCCAGTCCCTGMAAATACTACGGC 1417
 QY 913 tacggtgctggttaacccattggtccagctcaaggtggttcttcttcttcttcttcttcttct 972
 Db 1418 CATGGCGCAGGTAAACCGCTCGGCCGACCCAGGCGTCCGCTACGCTTAACAGACTCATC 1477
 QY 973 gctagatgaactcacttccagttcaagaacacacttctactaaccacacttcttcttcttcttct 1032
 Db 1478 GCCCGTCTCACCACCTCGCTGTCCAGATGACACACGCTCCACACACCTTGGACTCG 1537
 QY 1033 aaccagctacttcccatggaacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1092
 Db 1538 AACCCAGCTACCTTCCCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAACGGC 1597
 QY 1093 atgatatctatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1152
 Db 1598 ATCATCTCTATCTCTCTTCTTGGTCTGTATAACGGGACTAAAGCGCTGTCTACACAG 1657
 QY 1153 tctgtgaactcattgaagaactgacggttactcttcttcttcttcttcttcttcttcttcttcttct 1212
 Db 1658 ACCGTGGAAGATATACCCAGCAGATGGGTCTCTGCTCTGCTTGGACGCTTCCGTTTGTCT 1717
 QY 1213 gctagagcttactgtgaaatgatgcaatgcaagctgaaagaaacacacacacacacacacacacac 1272
 Db 1718 TCGGCTGTGTACGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1777
 QY 1273 ttggttaacgacagattgttccattgacggttcttcttcttcttcttcttcttcttcttcttcttct 1332
 Db 1778 TTGGTTAATGATCGCTGTGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1837
 QY 1333 aagagacgactcgttgaaggttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1392
 Db 1838 ACCCGGATAGCTTGTGAGGGGTTGAGCTTGTGATAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGT 1897
 QY 1393 tgttctgctta 1403
 Db 1898 TGTCTGCTTA 1908
 RESULT 12
 AR051916
 LOCUS 2363 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 7 from patent US 5830733.
 ACCESSION AR051916
 VERSION AR051916.1 GI:5975280
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2363)
 AUTHORS Nevalainen, H.K.M., Paloheimo, M.T., Miettinen-Oinonen, A.S.K.,
 Torkkeli, T.K., Cantrell, M., Piddington, C.S., Rambosek, J.A.,
 Turunen, M.K., Fagerstrom, R.B. and Houston, C.S.
 TITLE Nucleic acid molecules encoding phytase and pH2.5 acid phosphatase
 JOURNAL Patent: US 5830733-A 7 03-NOV-1998;
 FEATURES Location/Qualifiers
 source 1..2363
 BASE COUNT 559 a 732 c 510 g 562 t
 ORIGIN

Query Match 44.8%; Score 629.4; DB 9; Length 2363;
 Best Local Similarity 65.8%; Pred. No. 4.6e-160;

Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

QY	13	gtgtgtctactgttcattgccacattgttcttggtttccacatcccggtaccgcttgggtctct	72
DB	518	gtggcagctactgtatcgctgacaattctgtcagagatcacctccggactggcagctccccgcc	577
QY	73	cgttggaattctcactctgttgacactgttgacggttggttaccaatgtttcccaagaatt	132
DB	578	tgcgaanaatcaatccacttgggattacgggtcgatcaagggtatcaatgcttctccgagact	637
QY	133	tctcacttgtgggttacctactctccatactctcttcttggcagagaatctcgtattctt	192
DB	638	tgcgcattcttggggtcaatagcggcggttcttctctgscraacgaatcggccatctcc	697
QY	193	ccagacgttccagacgactgtagagttactttcgttccagtttctctagacacggtgct	252
DB	698	cctgatgtgccccgggttgcagagtcacttttccgtcaggtcctctcccgctcatggagcg	757
QY	253	agatacccaactctctcggtctcaaggttactctgtcttgattgaagactattccaaag	312
DB	758	cggtatccgacggagtccaaggcgagaanaatctcgcgtctcattgagagatccagcag	817
QY	313	aacgctactgcttoaagggttaagtcaggtttcttgaagacttacaacttcaactttgggt	372
DB	818	aacgtgacaccactttgatggaaaatattgcgttcttgagacatacaacttacagcttgggt	877
QY	373	gctgacgacttgctccactcgtgaaacccaaatgttaactctggtattaaagtctac	432
DB	878	gcagatgacctgactcctctcgagagcaggacctagttcaactccgggatcaaatgtttac	937
QY	433	agaagatacaaggcttgggtagaagattgttccattcattagagctctggtctcgac	492
DB	938	cagcgatatacgaatcggttcaagaagaacatcaitccgttcatccgatctcttggctcagc	997
QY	493	agagttattgctctcgtaaaaagttcattgaaggtttccaaactcgttaagttggtgac	552
DB	998	cgcgtgatcgccttcggcgagaaatttcaattgagggtcttccagagcaccagctgaagcat	1057
QY	553	ccaggttctcaaccacacaaagcttctcagttattaagcgtgatacttccagaagatcc	612
DB	1058	cctcgtgtcccagcggcgcaaatcgtcgcccaagatcgacgtgggtcatttccgagccagc	1117
QY	613	ggttcaacaacacttggaccacggtactgtactgctgtttcgaaagactcgaattaggt	672
DB	1118	tcattccaacaacactctcgaccacgacctgcactgtcttggaaagacagcgaattggcc	1177
QY	673	gacgagttgaagctaaactcactgcttctgttgctccagctattagagtagattgaa	732
DB	1178	gatacggtcgaaagccaaatttcaccgcccaggttgcgccccctccattcgtcaacgtctggag	1237
QY	733	gctgacttgcaggtgttaatttgactgacgagaagctgtgttacttgatggacatgtgt	792
DB	1238	aacgacctgtctggcggtgactctacacagacacagaagctgaccttacctcattggacatgtgc	1297
QY	793	ccattcgacacttgcgttagaaactctgacgtactgaaattgtctccatctcgtcttg	852
DB	1298	tccttcgacaccattccaccagcagctgcgacaccaagctgtcccccttctgtgacctg	1357
QY	853	ttaactcacgagaatgataccaatacagactacttgcagaagttgggttaagtactacgt	912
DB	1358	ttuaccccatgacgaatgatacttaccagcttaccagcttaccctcagcttgcacaccagctccaacacacactttggactct	1417
QY	913	tacggtgtcgtgtaacccattgggttcagctcaaggtgtggttcgctaagaattgatt	972
DB	1418	catggcgacaggttaaccgctcgcccgacccacggcggtcgcgcttaccgtatcaggctatc	1477
QY	973	gtagattgactcactctcaggtttaaagacacacttctactaacacacattttggactct	1032
DB	1478	gccccgtctacaccactctcccttgcacagatgacacaccagctccaacacacactttggactct	1537
QY	1033	aaccagctactttcccatggaacgtcactttgacgtgaactctctcacgacaact	1092
DB	1538	aaccacagctactttcccgctcaactctatctctacggcgacttttccacacatgacgc	1597

QY	1093	atgatactattttcttcgttttgggtttgtcaaacggtaccagccattgtctactact	1152		
Db	1598	ATCATCTATCATCTCTTGGCTTGGCTGTCTATCAACGGCACTAAGCCGCTGTCTACCACG	1657		
QY	1153	tctgttgaactctattgaagaacctgacggttactctgttcttcttgactgttccactcgt	1212		
Db	1658	ACCGTGGAGAAATATCACCAGACAGATGGGTCTCTGCTGCTGGACGGTTCGGTTGCT	1717		
QY	1213	gctagagcttaogttgaattgatgcaatgtcaagctgaaaaaggaacccattgttagagtt	1272		
Db	1718	TCGCGTCTACGTCGAGATGATGAGTCCAGTCCAGCGCGGACGAGGCCGTGTCGGTGC	1777		
QY	1273	tgtgttaacgacagagttgttccattgacagtgctgtgttgacaagttggttagatgt	1332		
Db	1778	TGTGTTAATGATCGGGGTTGTCGCCGTGATGGGTGTCCAATTGATGCTTTGGGGAGATG	1837		
QY	1333	aagagagacgacttcgttgaaggttcttcttcgctagatctggttgtaactggtctgaa	1392		
Db	1838	ACCCGGGATAGCTTGTGAGGGGTTGACCTTGTAGATCTGGGGGTGATTGGCGGAG	1897		
QY	1393	tgtttcgctta 1403			
Db	1898	TGTTCTGCTTA 1908			
RESULT 13					
AR053934					
LOCUS	AR053934	2379 bp	DNA		
DEFINITION	Sequence 1 from patent US 5834286.				
ACCESSION	AR053934				
VERSION	AR053934.1				
KEYWORDS	GI:5978796				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2379)				
TITLE	Nevalainen, H. K. M., Paloheimo, M. T., Pajunen, R. B.,				
	Miettinen-Oinonen, A. S. K., Turunen, M. K., Ramboisek, J. A.,				
	Piddington, C. S., Houston, C. S. and Cantrell, M. A.				
	Recombinant cells that express phytate degrading enzymes in desired ratios				
JOURNAL	Patent: US 5834286-A 1 10-NOV-1998;				
FEATURES	Location/Qualifiers				
source	1..2379				
BASE COUNT	561 a	735 c	518 g		
ORIGIN	565 t				
Query Match					
Best Local Similarity 44.8%; Score 629.4; DB 9; Length 2379;					
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;					
QY	13	gtcgtgctactgtccattgccaacttctggtttccacatcccggtacgccttgggtcct	72		
Db	534	GTGGGACTACTGATCGCTGACAAATCTGTGAGAGTCACTCCGGACTGGCAGTCCCGCC	593		
QY	73	cgttggttaattctcactcttgcactgttgacactgttgacggtgtttaccagtgttccagaatt	132		
Db	594	TCGAGAAATCAATCCACTTCGATACGGTTCGATCAAGGGTATCAATGCTTCTCCGAGACT	653		
QY	133	tctcactgtgggtgactactctcactctcactctctcttggcagacaaatcgtattct	192		
Db	654	TCGCATCTTTGGGTCAATACGCGCGCTTCTCTCTCTGCGCAACGAATCGGCCATCTCC	713		
QY	193	ccagacgttccagacactgtagattacttctcgttcaagtttcttagacagcgtgct	252		
Db	714	CCTGATGTGCCCGCGGTTCGAGATCACTTTCGCTCAGGTCTCTCCGTCATGAGCG	773		
QY	253	agatacccaactcttctcgtctaggcttactctgctttagtaagcattcctcgaag	312		
Db	774	CGTATCCGACCGAGTCCAAAGGCGCAAGAAATCTCGCTCTCATCAGGAGATCCAGCAG	833		

Qy	313	aacgctactgcttcaagggttaagtaacgcttcttgaagaacttaacacttctgggt	372
Db	834	AACGTACCACCTTTGATGTAATAATGCTCTCTGAAGACATACAACTACAGCTTGGGT	893
Qy	373	gctgaacactgactccatcgatgaacacccaataagtttaactctggtatataagttctac	432
Db	894	GCATGACCTGACTCCCTTCGGAGACAGAGAGTAGTCAACTCCGGCATCAAGTCTTAC	953
Qy	433	agaagatacaaggcttggcttagaagaattgttccattcattagagcttctggtctgac	492
Db	954	CAGCATACGAATCGCTCACAAAGAACATCATTCCTCGTTCATCCGATCTCTGGCTCAGC	1013
Qy	493	agagtattgtctgtgtaaaagttcattgaaggtttccaatctgtaagttggctgac	552
Db	1014	CGCGTGTGCTCGCTCCGGCGGAATTCATTGAGGGCTTCAGAGACCAAGCTCAAGGAT	1073
Qy	553	ccaggttctcaaccacacacgaagcttctccagttattaaagctgacattccagaagatcc	612
Db	1074	CCTGTGCCAGCGGGCCCAATCGTCCGCCAAGATCGACGTGGTTCATTTCCGAGGCCAGC	1133
Qy	613	ggttacaacaaactttggaccacggtactgttactgtcttcttgaagactctgaaatggt	672
Db	1134	TCATCCAAACACTCTCGACCCAGGACCTGCACTGTCTTTGAAGACAGCGAATTGGCC	1193
Qy	673	gacgacttgaagctaaactcaactcgtttgttcctccagctattagacttagtgaa	732
Db	1194	GATACCGTGAAGCGCAATTTCCACGCCACAGTTCGCCGCCCTCCATTCGTCAACGTCTGGAG	1253
Qy	733	gctgacttgcaggtgttactttgactgacgaagacgtgtgtttacttgatggacatggt	792
Db	1254	AACGACCTGTCTGGGTGACTCTCACAGACACAGAAGTGACCTACCTCATGGACATGCG	1313
Qy	793	ccattcgacactgtcgttagaactctgacgtactgaatgttctccattctgtgcttg	852
Db	1314	TCCTTCGACACCATCTCCACGACCGCTCGACACCAAGCTGTCCCGCTTCTGTGACCTG	1373
Qy	853	ttcacttcacgaagatgatccaatacgaactacttgaaagcttgggttaagtactacggt	912
Db	1374	TTACCCATGACGAATGATCCACTACGACTACCTCCAGTCCCTCCGTAATAATACTACGGC	1433
Qy	913	tacggtgtggttaacccattgggttcagctcaagtggttggtttcgttaacgaattgatt	972
Db	1434	CATGGCGCAGGTAAACCCGCTCGGCCCGACCCAGGCGCTCGCTAGCTTAACGAGCTATC	1493
Qy	973	gctagattgactcaactctccagttcaagacacacacttctactaaccacactttgactct	1032
Db	1494	GCCGCTCTCACCCACTCGCCTGTCCACGATGACACCACTCCAGTCCCTCCGTAATAATACTACGGC	1553
Qy	1033	aaccagactactttccattgaacgtactttgtacgtgacttctctcagcaacaact	1092
Db	1554	AACCCAGCTACCTTCGCGTCAACTCTACTCTCTACGCGGACTTTTCCACGATAACGGC	1613
Qy	1093	atgatatctatttcttctgcttgggtttgtaaacggttaccagcaattgtctactact	1152
Db	1614	ATCATCTCTATCCTCTTCTTGGCTCTGTACACGCGACTTAAGCCGCTGTCTACCAAG	1673
Qy	1153	tctgttgaactctatgaagaactcaggttactctgtcttctgactgttccattcgt	1212
Db	1674	ACCGTGGGAATATACCCAGACAGATGGGTCTCTGCTGTGGACGGTTCCTGTTGCT	1733
Qy	1213	gctagagcttacttgaatgatgaatgtcaagctgaaaggaacccattgggttagagtt	1272
Db	1734	TCGCGTCTGTACGTCGAGATGATGATGCCAGGCGGACGAGAGCGCGTGTCTCGTCTC	1793
Qy	1273	ttggttaacagacagatgtttccattgacggttctgtctgttgacaaagtgtgggttagatgt	1332
Db	1794	TTGGTTAATGATCGCGGTGTCCCGCTGCGATGGGTGTCGAATGATGCTTTGGGGAGATGT	1853
Qy	1333	aagagagacgacttctgtgaaggtttgtcttctcgtagatctggttgtaactcgggtgaa	1392
Db	1854	ACCCGGGATAGCTTTGTGAGGGGGTTGACCTTTTGCTAGATCTCTGGGCTGATTGGCGGAG	1913

Qy	1393	tgatttcgctta	1403
Db	1914	TGTTCTGCTTA	1924

RESULT 14

ASNPHYTAS 2379 bp DNA PLN 23-MAR-1994

LOCUS Aspergillus niger var awamori phytase gene, complete cds.

DEFINITION L02421

ACCESSION L02421

VERSION L02421.1

KEYWORDS GI:166518

SOURCE phytase.

ORGANISM Aspergillus niger (strain ALK0243, sub-species awamori) DNA.

REFERENCE 1 (sites)

AUTHORS Aspergillus niger

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

TITLE Piddington, C.S., Houston, C.S., Paloheimo, M., Cantrell, M., Miettinen-Oinonen, A., Nevalainen, H. and Rambosek, J. The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var. awamori

JOURNAL Gene 133 (1), 55-62 (1993)

MEDLINE 94040796

REFERENCE 2 (bases 1 to 2379)

AUTHORS Carter, J.R., Franden, M.A., Aebersold, R.H. and McHenry, C.S. Molecular cloning, sequencing and overexpression of the gene encoding the psi subunit of E. coli DNA polymerase III holoenzyme Unpublished (1992)

JOURNAL Location/Qualifiers

FEATURES

source 1..2379

organism="Aspergillus niger"

/db_xref="taxon:5061"

420..463

/citation=[2]

/evidence=experimental

join(420..463,566..1925)

/EC_number="3.1.3.8"

/citation=[2]

/codon_start=1

/function="conversion of phytate to inositol and inorganic phosphorus"

/evidence=experimental

/product="phytase"

/protein_id="AA16898.1"

/db_xref="GI:166519"

/translation="MGVSAVLLPLVLLAGVTSGLVAPSRNOSTCDTVDDQYQCFSET SHLWGYAPFFSLANESAIISPDVAGCRVTAQLVSRHGARYPTESKGYKYSALIEEI QONVTFDGKYAFKATYNSLGADDLTPFGEOLVNSGKIFQYVESLIRNIIPIRS SSSRVIASGEKFIQFQSTKLKDPRAOPGSSPKIDVVISSSSNNTLDPGCTVF EDSEADTVEANFTAFAPSIQRLENLISGVTLDTEVTYLMDCSFDTLSTSPVD KISPFCDLEHDEWIHYDLQSKKYYHGAGNPLGPTGGVGYANELARTLHSPVHD DTSNNHTLDSNPATPLNSTLYADFSDHNGIISILFALGLYNGTAPLSTIVENITQT DGFSSAWTVPFASRLYVENMQCAEQEPLRVLRVNDVRVPLVPLHGPCIDALGRCTRDSFV RGLSPARSGGDNAECSA"

566..1925

/note="putative"

/citation=[2]

518 g 565 t

BASE COUNT 561 a 735 c

ORIGIN

Query Match 44.8%; Score 629.4; DB 13; Length 2379;

Best Local Similarity 65.8%; Pred. No. 4.6e-160;

Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

Qy	13	gtcgtgctactgtccattgcacattgttcggtttccacattccggtaccgcttgggtcct	72
Db	534	GTGGGACTACTGATCGTGACAAATCTGTGCAGAGTCACTCCGGAATGGCAGTCCCGCC	593
Qy	73	cgtgtgtaattctactcttctgtgacactgttgacggtgtttaccaatgtttccagaaatt	132
Db	594	TCGAGAAATCAATCACTTGGATACGGTACGATCAAGGGTATCAATGCTTCTCCGAGCT	653

QY 133 tctcactgttggttacactctccatactctctcttcttggcagacgaatctgtctattct 192
 Db 654 TCGCATCTTTGGGGTCAATACGCGCGTCTCTCTCTGGCAACGAATCGCCATCTCC 713
 QY 193 ccagacttcagacgactgtagattacttctgctcaagtcttctgtagacacgggtgct 252
 Db 714 CCGTATGTCGCGCGGTTGCAGAGTCACTTTCGCTAGGTCCTCTCCCGTCATGGAGCG 773
 QY 253 agataccaaactctctgctcctaaggtctactctctgcttctgattgaagctattcaaaa 312
 Db 774 CGGTATCCGACCGAGTCCCAAGGGAAGAAATACTCCGCTCTCATTTAGGAGATCCACGAG 833
 QY 313 aacgctactgcttccaagggtaagtagcgtttcttgaagacttaacaactcaacttgggt 372
 Db 834 AAGGTGACCACTTTGATGGAATATGCCCTTCTGAAGACATACAACATACAGCTGGGT 893
 QY 373 gctgacgactgactcattcctggtgaacacaaatggttaactctggtatttaagttctac 432
 Db 894 GCAGATGACCTGACTCTCGGAGAGCAGGAGCTAGCTAACTCCGGCATCAAGTTCTAC 953
 QY 433 agaagatacaaggcttgggttagaagaattgttccattcattagagcttctggtttgac 492
 Db 954 CAGCGATACGATCGCTCAAGGAAGAACATCTCCGTTCTATCGATCCTCTGGCTCCAGC 1013
 QY 493 agagtattgcttctgctgaaggttctcattgaaggttccattcctgcttaagttggctgac 552
 Db 1014 CGCGTGATCGGCTCCGCGGAGAAATTCATTGAGGGTTCAGAGACACCAAGCTGAAGGAT 1073
 QY 553 ccaggttctcaac 612
 Db 1074 CCGTGGCCGACCGCGGCCACTCGTCCGCAAGATCGACGCTGATTTCCGGGCCAGC 1133
 QY 613 ggttac 672
 Db 1134 TCATCCAAACACTCTCGACCCAGGACCTGCACCTGCTTTTGAAGACGACGAATGGCC 1193
 QY 673 gacgacttgaagacttaactcactgcttctgctccagctattagagctagattgaa 732
 Db 1194 GATACCGTCGAAGCAATTTTCACCGCACGCTCGGCCCTCCATTCTCAACGCTCGGAG 1253
 QY 733 gctgactgctcaggtgttacttgaactgacgaagacgttgttactgtatgacatgtct 792
 Db 1254 AACGACCTGCTCGCGTGACTCTCAGACACACAGAAAGTGAACCTACCTCATGACATGTC 1313
 QY 793 ccactgacactgctgtagaactctgacgctactgactgaattgtctccattctgtgcttg 852
 Db 1314 TCCTTGACACCACTCTCCACGACCGCTGACACACCAAGCTGTCCCTCTCTGTGACCTG 1373
 QY 853 ttcaactcagcagaatggatccaatcagactacttgcgaagcttgggttaagtactacggt 912
 Db 1374 TTCACCATGACGAATGGATPCACTACGACTACCTCCAGTCCCTGAAAAAATACTACGGC 1433
 QY 913 tacggtgctggttaacccattggttcacgctcaggtgctggttctgctcaacgaattgatt 972
 Db 1434 CATGGCGCAGGTAACCGCTCGCGCCGACCCAGGCGGTGGGTACGCTAACGAGCTCATC 1493
 QY 973 gctagattgactcactctccagltcaagacacacacttctactaaccacacacttggactct 1032
 Db 1494 GCGCGTCTCACCCACTCGGCTCTCCACGATGACACGAGCTCCCAACACACACCTTGACTCG 1553
 QY 1033 aacccagctacttcccaactgacgctacttgttacgtgacttctctcaagacacact 1092
 Db 1554 AACCCAGCTACCTTCCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAAACGGC 1613
 QY 1093 atgatctatttcttctgcttgggttgggttgaacacggtacacagcattgtctactact 1152
 Db 1614 ATCATCTCTACTCTTTGGCTTGTACACGCGCATTAAGCCGCTGTCTACACAG 1673
 QY 1153 tctgttgactctattgaagaactgaaggttactctgcttcttctgactgttcccttgcct 1212
 Db 1674 ACCGTGGAGAAATATACCCACGACAGATGGGTCTCTGCTGCTTGGACGGTTCCGCTTGCT 1733

QY 1213 gctagagcttacgttgaatgatgcaatgtcaagctgaaaggaacacattggttagagtt 1272
 Db 1734 TCCGCTGTACGTGAGATGTCAGATGCCAGCGCAGCAGCGCTGCTCGGTTC 1793
 QY 1273 ttggttaacagacagagttgttccattgcacggttgcgttgcattgacaagtgggttagatgt 1332
 Db 1794 TTGTTTAATGATCCGCTGTCCCGCTGCATGGGTGTCATGTCATGATGCTTTGGGAGATGT 1853
 QY 1333 aagagagacgactcttgaaggttgcattcttctgcttagatctcgttgtaactgggtgaa 1392
 Db 1854 ACCCGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGCGGAG 1913
 QY 1393 tgttctgctta 1403
 Db 1914 TGTTCTGCTTA 1924

RESULT 15
 AB022700
 LOCUS
 DEFINITION
 Aspergillus niger gene for phytase, complete cds.
 ACCESSION
 AB022700
 VERSION
 AB022700.1 GI:4185609
 KEYWORDS
 phytase.
 SOURCE
 Aspergillus niger (strain:SK-57) DNA.
 ORGANISM
 Aspergillus niger
 Aspergillus niger
 Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE
 1 (bases 1 to 1515)
 Nagashima,T., Kondo,H., Anazawa,H. and Terasaki,Y.
 Phytase having high-affinity for phytic acid
 Published only in DataBase (1999) In press
 2 (bases 1 to 1515)
 Nagashima,T., Kondo,H., Anazawa,H. and Terasaki,Y.
 Direct Submission
 Submitted (21-JAN-1999) to the DDBJ/EMBL/GenBank databases. Tadashi
 Nagashima, Shin Nihon Chemical Co.,Ltd, Research and Development
 Div., Showa-cho 19-10, Anjo, Aichi 446-0063, Japan
 (E-mail:nagashima@dl.dion.ne.jp, Tel:81-566-76-5171(ex.264),
 Fax:81-566-75-0010)

FEATURES
 Location/Qualifiers
 1..1515
 /organism="Aspergillus niger"
 /strain="SK-57"
 /db_xref="taxon:5061"
 join(1..44,156..1515)
 /gene="phyA"
 join(1..44,156..1515)
 /gene="phyA"
 /codon_start=1
 /product="phytase"
 /protein_id="BAA74433.1"
 /db_xref="GI:4185610"
 /translation="MGVSALLPLVLLSLVPSRQSTGLAVPSRQSTCDTVQGYQCRSET
 SHLWGOVAPFSLANKSAISPDVPCGHVTFQVLSRHGARIPTDSKKAKYALIEEI
 QNATTEKYAFLEKTYNTSLGADLTPEGLVNSGVKVFYRSESTRTNTPVTRIS
 SCSRRVIAENKFIQGTSLKLRPAQSPKSPKIDVVISEASTNNTLDPTCTVF
 EDELADDEANFTATTFVSIQRLENDLSGLDTEVTYLMDCSFTIISTVDT
 KUSPFDLFTHEWINDYLIQSLNKYHGHGAGNPLCPTOGVGYANELIARLTHSPVHD
 TYSNNHTLSDNPATFPLNSTLVDADFSDHNGIISILFALGLYNGTKPLSTTAENITQT
 DGFSSAWTVPFASRMVYEMMOCQSEQEPILRVILVNDRVVPLHGGCPVDALGRCTRSFV
 KGLSPARSGGDWCECA"

BASE COUNT 317 a 478 c 365 g 355 t
 ORIGIN

Query Match 43.6%; Score 612.4; DB 12; Length 1515;
 Best Local Similarity 65.5%; Pred. No. 1.8e-155;
 Matches 895; Conservative 0; Mismatches 471; Indels 0; Gaps 0;

QY 38 tgtctggttccacatccggtacgccttgggtctgctggttaattctcactcttgaca 97
 Db 149 TGCACAGAGTCACTCCGGACTGGCAGTCCCGCCCTCGAGAAATCAATCCACTTGGGATA 208

QY	98	ctgttgacggtggttaacaaatgtttccacagaataatttcacttggtggtggtacactactctc	157
Db	209	CGGTGATCAGGGGTATCAATGCTTCGGAGACTTCGGCATCTTTGGGGCCAATACGCGC	268
QY	158	catactctcttggcagacgaatctcgtatttctcagacagcttccacagacagctgtagag	217
Db	269	CGTCTTTCTCTGSCAACAAATCGGCCATCTCCCTGATGTTCTTCGCGGATGCCCATG	328
QY	218	ttactttcgttcaagtgttctagacacagtgctagatacccaactcttctcgtctcta	277
Db	329	TCACCTTTTCGCCAGGTTCTTCGCCCATGAGACAGCATCCACGACTCCCAAGGSCA	388
QY	278	agcgttactctgttggattgaagctattcaaaagacgctactgctttcaagggttaagt	337
Db	389	AGAAATACTCCGCTCTCATCGAGGAGATCCAGCAGACGCGACAACTTCGAGGGGAAT	448
QY	338	acgctttcttgaaacttaacactacactttgggtgctgacgacttgactccattcgtg	397
Db	449	ATGCCCTTCGTAAGACATACAACTACAGCCTGGCGCGGATGATCTGACTCCCTTCGGAG	508
QY	398	aaacacaaatgcttaactctggtattaaagtctctacagaagatacaaggctttggtctaga	457
Db	509	ACAGAGAGCTGTCAACTCCGCGCTCAAGTCTACACGCGATCGAATCGCTCACAGAA	568
QY	458	agattgttccattcattagagcttctggttcttgacagagttattgcttctcgtctgaaaa	517
Db	569	ACATTGTCCGTTCCGATCCCTCAGGCTCCAGCGGCTGATTTGCCCTTCGGCAATAAAT	628
QY	518	tcattgaaggtttccaactctgtaagtgtggtgacccaggttctcaaccacacaaagctt	577
Db	629	TCATCGAGGGCTTCACAGACCTAAAGTGAAGGATCCCTGTCGCCCGCGGCCAATCGT	688
QY	578	ctccagttattacgtgatactccacagaggtccggtttacacaaacactttggaccacy	637
Db	689	CGCCCAAGATCGAGTGGTCAATTCAGAGGCCAGCACATCCAAACACACTCTCGATCCGG	748
QY	638	gtacttgactgcttccgaagactctgaattaggtgaagacgttgaagacttaacttcaact	697
Db	749	GCACCTGCACCGTTTTCGAAGATAGCGAATTGGCGGATGACATCGAAGCCAAATTCACCG	808
QY	698	cttgtgctccagctattagactagactgaagctgactgacgtgacgtgttactttga	757
Db	809	CCACGTTCCGTCCTCCATTCGTCACAGCTCGAGAACACTTTGCTGCGGTGCTCTCA	868
QY	758	ctgacgaagacggtgttacttgatggacatgtgtccattcgcacactgctgctagaactt	817
Db	869	CGGACACAGAAGTGACCTACCTCATGGACATGTGCTCTCGACACCATCTCCACCAGCA	928
QY	818	ctgacgctactgaattgcttccacttctgtgtgttcttactcagcagcaatggtccaat	877
Db	929	CCGTCGACACCAAGCTGCCCCCTTGTGACCTGTTCACCCCATGAAGAATGATCAACT	988
QY	878	acgactacttgcaagcttgggttaagtactacggttacggtgctggttaacccattgggtc	937
Db	989	ACGACTACCTCCAGTCCCTGAAACAATACTACGCCCATGGCCAGTACCGCGCTCGGCC	1048
QY	938	cagctcaaggtgtgttgcgttaacgaattgattgtagattgactcaactctccagttc	997
Db	1049	CGACCCAGGGCGTCGGCTACGCTAACGAGCTCATCGCCGCTCTCACCCACTCGCCCTGCC	1108
QY	998	aagaccacacttctactaaccacacttgacttaacctaacccagctactttccattgaagc	1057
Db	1109	ACGATGACACAGCTCCAAACACACATTTGACTCCAAACCCGGCTACTTTCCCGCTCACT	1168
QY	1058	ctactttgacgctgacttctcagcacaacacatgatatttttcttcgctttgg	1117
Db	1169	CCACTCTATGCGGACTTTTCGATGATAACGGCATCATCTCTATCTCTTTGCTTTGG	1228
QY	1118	gtttgtaacaggttaccacagccattgtctactactctctgtgaatctattgaagaactg	1177
Db	1229	GTCTGTACACAGCGGCCAAGCGCGTGTCTTCCACGACCGGAGAAATATACCCAGACCG	1288

Search completed: October 26, 2001, 17:47:03
Job time: 8779 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 17:50:17 ; Search time 210.38 seconds
(without alignments)

Title: US-09-488-265-28

Perfect score:
Sequence.

sequence. I atgggcgtgttcgtcgtgct.....gggctgaatgtttcgcttaa 1404

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*

1:	/SIDS1/cgcdata/geneseq/geneseqn/NA1980.DAT.*
2:	/SIDS1/cgcdata/geneseq/geneseqn/NA1981.DAT.*
3:	/SIDS1/cgcdata/geneseq/geneseqn/NA1982.DAT.*
4:	/SIDS1/cgcdata/geneseq/geneseqn/NA1983.DAT.*
5:	/SIDS1/cgcdata/geneseq/geneseqn/NA1984.DAT.*
6:	/SIDS1/cgcdata/geneseq/geneseqn/NA1985.DAT.*
7:	/SIDS1/cgcdata/geneseq/geneseqn/NA1986.DAT.*
8:	/SIDS1/cgcdata/geneseq/geneseqn/NA1987.DAT.*
9:	/SIDS1/cgcdata/geneseq/geneseqn/NA1988.DAT.*
10:	/SIDS1/cgcdata/geneseq/geneseqn/NA1989.DAT.*
11:	/SIDS1/cgcdata/geneseq/geneseqn/NA1990.DAT.*
12:	/SIDS1/cgcdata/geneseq/geneseqn/NA1991.DAT.*
13:	/SIDS1/cgcdata/geneseq/geneseqn/NA1992.DAT.*
14:	/SIDS1/cgcdata/geneseq/geneseqn/NA1993.DAT.*
15:	/SIDS1/cgcdata/geneseq/geneseqn/NA1994.DAT.*
16:	/SIDS1/cgcdata/geneseq/geneseqn/NA1995.DAT.*
17:	/SIDS1/cgcdata/geneseq/geneseqn/NA1996.DAT.*
18:	/SIDS1/cgcdata/geneseq/geneseqn/NA1997.DAT.*
19:	/SIDS1/cgcdata/geneseq/geneseqn/NA1998.DAT.*
20:	/SIDS1/cgcdata/geneseq/geneseqn/NA1999.DAT.*
21:	/SIDS1/cgcdata/geneseq/geneseqn/NA2000.DAT.*
22:	/SIDS1/cgcdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1404	100.0	1404	20	AAZ31520	Consensus phytase-
2	1404	100.0	1404	21	AAZ73233	Consensus phytase
3	1404	100.0	1404	21	AAZ59715	DNA encoding a mut
4	1388	98.9	1404	21	AAZ73291	Consensus phytase
5	1384.8	98.6	1404	21	AAZ73290	Consensus phytase
6	1360.8	96.9	1426	20	AAZ27423	Ascomycete consens
7	1360.8	96.9	1426	20	AAZ23022	Fungal phytase gen
8	1360.8	96.9	1426	21	AAZ73231	Consensus phytase
9	1360.8	96.9	1426	21	AAZ59637	Consensus phytase-
10	1327.2	94.5	1404	20	AAZ31521	Consensus phytase
11	1327.2	94.5	1404	21	AAZ73234	Consensus phytase

12	1327.2	94.5	1404	21	AAZ59716	DNA encoding a mut
13	1324	94.3	1404	21	AAAY3293	Consensus phytase
14	1320.8	94.1	1404	21	AAAY3292	Consensus phytase
15	1317.6	93.8	1426	21	AAAY3232	Consensus phytase
16	1317.6	93.8	1426	21	AAZ59642	DNA encoding phyta
17	1292	92.0	1426	20	AAZ31523	Consensus phytase-
18	1292	92.0	1426	21	AAAY3236	Consensus phytase
19	1292	92.0	1426	21	AAZ59738	DNA encoding phyta
20	636.6	45.3	1404	18	AAQ61136	Consensus phytase
21	635	45.2	1404	12	AAQ11175	Aspergillus ficum
22	635	45.2	1404	20	AAZ27421	Chromosomal phytas
23	631.8	45.0	1404	12	AAQ13878	A. ficum phytase
24	631	44.9	6756	12	AAQ11174	Phytase gene. Asp
25	631	44.9	6756	18	AAQ65137	Sequence, from ove
26	629.4	44.8	1404	20	AAZ31522	Aspergillus ficum
27	629.4	44.8	1404	21	AAAY3235	A. fumigatus phyta
28	629.4	44.8	1404	21	AAAY3235	Aspergillus fumiga
29	629.4	44.8	2363	15	AAZ59717	DNA encoding a mut
30	629.4	44.8	2379	15	AAO58126	Phytase gene. Asp
31	604.4	43.0	1515	18	AAO56914	A. niger phytase g
32	581.2	41.4	1571	19	AAO33144	Aspergillus niger
33	581.2	41.4	1571	20	AAZ27432	Aspergillus fumiga
34	578.4	41.2	1931	19	AAO33142	A. fumigatus phyta
35	578.4	41.2	1931	20	AAZ27413	Aspergillus nidula
36	559.6	39.9	1912	17	AAO37473	A. nidulans phytas
37	549	39.1	1845	19	AAO33143	Phytase chimeric g
38	547.4	39.0	1845	20	AAZ27416	Phytase chimeric g
39	539.8	38.4	1567	19	AAO33145	Talaromyces thermo
40	539.8	38.4	2327	17	AAO37336	T. thermophilus ph
41	539.8	38.4	2327	17	AAO37336	Aspergillus terreu
42	501	35.7	1932	22	AAZ27414	Phytase gene. Asp
43	500.4	35.6	1454	22	AAZ77028	A. terreus phytase
44	500.4	35.6	1584	22	AAZ77030	P.hordei phytase c
45	407.8	29.0	4898	22	AAZ77029	P. hordei phytase c
45	407.8	29.0	4898	22	AAZ77031	P. hordei phytase c

ALIGNMENTS

RESULT 1

AAZ31520
ID AAZ31520 standard; DNA; 1404 BP.

AAZ31520;

06-JAN-2000 (first entry)

Consensus phytase-1-thermo(8)-Q50T-K91A coding sequence.

Phytase; animal feed preparation; thermostable phytase; transgenic plant;
KW
consensus sequence; ds.
KW

OS Synthetic.

XX

PN ~~WO948380-A1~~

XX
PD 30-SEP-1999

FD
XX
30-SEP-1999.

PF 22-MAR-1999;

XX
XX

PR 23-MAR-1998;
PR 19-JUN-1998;

PR 18-SEP-1998;
PR 19-JUN-1998;

22-JAN-1999;

22-JAN-1999;

XX
XX

Downloaded from ascelibrary.org by New York University on 06/01/15. Copyright ASCE, all rights reserved.

XX
2A } NOVO-
(NOVO)

Petersen S;

XX

OR WPI; 1999-591

OR P-PSDB; AAY43

PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 9; 7lpp; English.

XX This sequence encodes the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.
 XX Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;

Query Match 100.0%; Score 1404; DB 20; Length 1404;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcggttctgctgctactgtccattgccaacctgttgggtccacatccgggtacc 60
 Db 1 atggcggttctgctgctactgtccattgccaacctgttgggtccacatccgggtacc 60
 QY 61 gcoettgggtctctggttaattctcaactttgtgacactgttgacggtgggtaccatgt 120
 Db 61 gcoettgggtctctggttaattctcaactttgtgacactgttgacggtgggtaccatgt 120
 QY 121 ttccagaattctcaactgtgggtacacttctccatacttcttggcagacgaa 180
 Db 121 ttccagaattctcaactgtgggtacacttctccatacttcttggcagacgaa 180
 QY 181 tctgctatttctccagacgttccagacgactgtagattacttcttcaagtgtgtct 240
 Db 181 tctgctatttctccagacgttccagacgactgtagattacttcttcaagtgtgtct 240
 QY 241 agacacggtgctagataccacacttcttctgctctaaagcttactctgttgattgaa 300
 Db 241 agacacggtgctagataccacacttcttctgctctaaagcttactctgttgattgaa 300
 QY 301 gctattcaaaaagacgctactgttcttcaagggttaagtacgcttcttgaagacttacac 360
 Db 301 gctattcaaaaagacgctactgttcttcaagggttaagtacgcttcttgaagacttacac 360
 QY 361 tacacttgggtgctgacgactgactcattctggtgaaacacaaatggttaactctgt 420
 Db 361 tacacttgggtgctgacgactgactcattctggtgaaacacaaatggttaactctgt 420
 QY 421 attaaagttcacagaagatacaaggcttctggtctagaagaattgttccatttagagct 480
 Db 421 attaaagttcacagaagatacaaggcttctggtctagaagaattgttccatttagagct 480
 QY 481 tctggttctgacagattattgtctctgctgaaagtctcattgaagggttccaaatctgt 540
 Db 481 tctggttctgacagattattgtctctgctgaaagtctcattgaagggttccaaatctgt 540
 QY 541 aagtgtgctgaccaggttctcaacacacacacacacacacacacacacacacacacacac 600
 Db 541 aagtgtgctgaccaggttctcaacacacacacacacacacacacacacacacacacacac 600
 QY 601 ccagaagatccgggttac 660
 Db 601 ccagaagatccgggttac 660
 QY 661 tctgaattaggtgacgacggttgaagctaaacttcaactgttctgctccagctattaga 720
 Db 661 tctgaattaggtgacgacggttgaagctaaacttcaactgttctgctccagctattaga 720
 QY 721 gctagattgaagctgacttgcaggtgttacttgaactaacacacacacacacacacacacac 780
 Db 721 gctagattgaagctgacttgcaggtgttacttgaactaacacacacacacacacacacacac 780

QY 781 atggacatgtgtcattcagacacactgtcgtacgaacttctgacgctaciacgaattgtctcca 840
 Db 781 atggacatgtgtcattcagacacactgtcgtacgaacttctgacgctaciacgaattgtctcca 840
 QY 841 tctgtgcttgttctcactcagcagcaatggatccaatacgaactacttgcgaagcttgggt 900
 Db 841 tctgtgcttgttctcactcagcagcaatggatccaatacgaactacttgcgaagcttgggt 900
 QY 901 aagtaactacggttacggtgctggttaacccattgggtccagctcaaggtgtgttctcgt 960
 Db 901 aagtaactacggttacggtgctggttaacccattgggtccagctcaaggtgtgttctcgt 960
 QY 961 aacgaattgattgctagattgactcactctccagttccaagacacacacttctactaacac 1020
 Db 961 aacgaattgattgctagattgactcactctccagttccaagacacacacttctactaacac 1020
 QY 1021 acttggacttaacccagctacttccattgaaacgctacttctgacgctgacttctct 1080
 Db 1021 acttggacttaacccagctacttccattgaaacgctacttctgacgctgacttctct 1080
 QY 1081 cagacacacactatgatactatttctcgttctggtttgtacaacggtaccacagcca 1140
 Db 1081 cagacacacactatgatactatttctcgttctggtttgtacaacggtaccacagcca 1140
 QY 1141 ttgtactacttctgttgatctatttgaagaactgaacggttactctgtcttcttggact 1200
 Db 1141 ttgtactacttctgttgatctatttgaagaactgaacggttactctgtcttcttggact 1200
 QY 1201 gttccattcgtcgtagagcttccgttgaatgatcaaatgtcaacgtgaaaggaacca 1260
 Db 1201 gttccattcgtcgtagagcttccgttgaatgatcaaatgtcaacgtgaaaggaacca 1260
 QY 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttgtgctgttgacaag 1320
 Db 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttgtgctgttgacaag 1320
 QY 1321 ttggttagagttgaagagacgacttccgttgaagggttcttcttcgtagatctggttgg 1380
 Db 1321 ttggttagagttgaagagacgacttccgttgaagggttcttcttcgtagatctggttgg 1380
 QY 1381 aactgggtgactgttctcgttaa 1404
 Db 1381 aactgggtgactgttctcgttaa 1404

RESULT 2

AAA73233

ID AAA73233 standard; DNA; 1404 BP.

AC AAA73233;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 1 thermo 8 q50t, k91a polynucleotide SEQ ID NO:28.
 DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 XX temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure; ds.

OS Synthetic.

XX WO200043503-A1.

PN 27-JUL-2000.

PD 21-JAN-2000; 2000WO-DK000025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX
PI
XX
DR
DR
XX
PT
PT
PT
PS
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SO

Lehmann M;
WPI: 2000-491161/43.
P-PSDB; AAB20526.

Novel phytases with improved properties such as temperature stability, pH stability and substrate specificity, for use in pharmaceuticals and compound foods and feeds -

Claim 10; Fig 7a-c; 240pp; English.

The present invention describes improved phytases, preferably with increased thermostability, and methods for producing them. The methods can be used for producing phytases with improved properties e.g. temperature stability, pH stability, pH profile, temperature profile, specific activity, substrate specificity, substrate cleavage pattern, substrate binding, position specificity, the velocity and level of release of phosphate from corn, reaction rate, phytate degradation rate, and end level of released phosphate. The phytases can be used to produce pharmaceutical compositions or compound food or feeds. The feed can be used to reduce levels of phytate in animal manure, by converting it into lower inositol phosphates and/or inositol and inorganic phosphate. The present sequence encodes a phytase sequence from the present invention.

Sequence 1404 BP: 329 A; 319 C; 304 G; 452 T; 0 other;

Query Match 100.0%; Score 1404; DB 21; Length 1404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggggcgttcgtcgctactgtccattgccacctgttgcgttccacatccgtacc 60
DB 1 atggggcgttcgtcgctactgtccattgccacctgttgcgttccacatccgtacc 60

QY 61 gcttgggtcctcgtgtaattctcactctgtgacactgttgacgttggttaccatgt 120
DB 61 gcttgggtcctcgtgtaattctcactctgtgacactgttgacgttggttaccatgt 120

QY 121 ttccagaaattctcactgtggttgacactctccactcttccactctcttggcagacgaa 180
DB 121 ttccagaaattctcactgtggttgacactctccactcttccactctcttggcagacgaa 180

QY 181 tctgctattctccagactgtccagacgactgtagacttacttctgcttcaagtgtct 240
DB 181 tctgctattctccagactgtccagacgactgtagacttacttctgcttcaagtgtct 240

QY 241 agacacggtgctagataccacactctctcgtctaaagcttactctcgttggattgaa 300
DB 241 agacacggtgctagataccacactctctcgtctaaagcttactctcgttggattgaa 300

QY 301 gctattccaaagacgctactgcttccaaagggttaagctgcttcttgaagacttacac 360
DB 301 gctattccaaagacgctactgcttccaaagggttaagctgcttcttgaagacttacac 360

QY 361 tacatttgggtgctcagactgactcattcgttgcctgaaacaaatgttaactctggt 420
DB 361 tacatttgggtgctcagactgactcattcgttgcctgaaacaaatgttaactctggt 420

QY 421 attaatctcagaagataccaggttggctgagaagattgttccattcattagact 480
DB 421 attaatctcagaagataccaggttggctgagaagattgttccattcattagact 480

QY 481 tctggttctgacagagtattgtctctgctgaaagttcattgaaggtttccaactctgt 540
DB 481 tctggttctgacagagtattgtctctgctgaaagttcattgaaggtttccaactctgt 540

QY 541 aagttggctgacccaggttctcaaccacacacacacacacacacacacacacacacac 600
DB 541 aagttggctgacccaggttctcaaccacacacacacacacacacacacacacacacac 600

QY 601 ccagaaggtatccggtttacaaacacactttggaccacggttactgttactgctttcgaagac 660
DB 601 ccagaaggtatccggtttacaaacacactttggaccacggttactgttactgctttcgaagac 660

QY 661 tctgaattagtagtgcagacgttgaaacttaacttaacttctgcttctgctccagctattaga 720
DB 661 tctgaattagtagtgcagacgttgaaacttaacttaacttctgcttctgctccagctattaga 720

QY 721 gctagattggaagctgacttgcaggtgttactttgactgacgaaagacattgttacttg 780
DB 721 gctagattggaagctgacttgcaggtgttactttgactgacgaaagacattgttacttg 780

QY 781 atggacatgtgtccattcgcacactgtcgtgataaacttctgacgctactgaattgtctcca 840
DB 781 atggacatgtgtccattcgcacactgtcgtgataaacttctgacgctactgaattgtctcca 840

QY 841 tctgtgcttcttctcactcagcaaatgataccaaacttaacttctgacgctactgaattgtctcca 900
DB 841 tctgtgcttcttctcactcagcaaatgataccaaacttaacttctgacgctactgaattgtctcca 900

QY 901 agtactacggttactgctggtgtgtaaccattgggttccagctcaaggtgtgtgttctgct 960
DB 901 agtactacggttactgctggtgtgtaaccattgggttccagctcaaggtgtgtgttctgct 960

QY 961 aacgaattgattgctagattgactcactctccagttccaaagcacttctactaaccac 1020
DB 961 aacgaattgattgctagattgactcactctccagttccaaagcacttctactaaccac 1020

QY 1021 acttggactctaaacacgactacttccattgacgctacttctgacgctacttctct 1080
DB 1021 acttggactctaaacacgactacttccattgacgctacttctgacgctacttctct 1080

QY 1081 caccagacacactatgatattcttctgcttctggtttgtacacggttaccagcga 1140
DB 1081 caccagacacactatgatattcttctgcttctggtttgtacacggttaccagcga 1140

QY 1141 ttgtctactactctctgtaactctattgaaacacgctgacttctcttctggtact 1200
DB 1141 ttgtctactactctctgtaactctattgaaacacgctgacttctcttctggtact 1200

QY 1201 gttccattcgtctgtagagcttactgtgaaatgagcaatgacgtgaaagaaacga 1260
DB 1201 gttccattcgtctgtagagcttactgtgaaatgagcaatgacgtgaaagaaacga 1260

QY 1261 ttggttagagtttgggttaacgacagagttgttccattgcacggttgcgtgtgacaag 1320
DB 1261 ttggttagagtttgggttaacgacagagttgttccattgcacggttgcgtgtgacaag 1320

QY 1321 ttgggttagatgttaagagagacgacttctgtaaggtttcttctgctagatctggtggt 1380
DB 1321 ttgggttagatgttaagagagacgacttctgtaaggtttcttctgctagatctggtggt 1380

QY 1381 aactgggctgaattgttctgcttaa 1404
DB 1381 aactgggctgaattgttctgcttaa 1404

RESULT 3

AAZ59715
ID AAZ59715 standard; DNA; 1404 BP.
XX AC AAZ59715;
XX AC AAZ59715;
DT 19-APR-2000 (first entry)
XX DNA encoding a mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.
XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
KW thermostable; animal feed; monogastric animal; phytate phosphorus;
KW phosphate availability; consensus; mutant; ds.
XX OS Aspergillus terreus 9A1.
OS Aspergillus terreus cbs16.46.

OS	Aspergillus niger var. awamori.	61	gcttgggtcctcggtgtaattctcactcttctgacactgttgacggtggttaccatgt	120
OS	Aspergillus niger str. NRRL3135.	Db		
OS	Aspergillus fumigatus ATCC13073.	61	gcttgggtcctcggtgtaattctcactcttctgacactgttgacggtggttaccatgt	120
OS	Aspergillus fumigatus ATCC32722.	QY		
OS	Aspergillus fumigatus ATCC58128.	121	ttccagaaatttctcactcttctgacactgttgacactcttcttcttggcagacgaa	180
OS	Aspergillus fumigatus ATCC26906.	Db		
OS	Aspergillus fumigatus ATCC32239.	121	ttccagaaatttctcactcttctgacactgttgacactcttcttcttggcagacgaa	180
OS	Emericella nidulans.	QY		
OS	Talaromyces thermophilus ATCC20186.	181	tctgtatttctccagactgtccagagactgttagagttacttctggttcaagtgttct	240
OS	Myceliophthora thermophila.	Db		
OS	Synthetic.	181	tctgtatttctccagactgtccagagactgttagagttacttctggttcaagtgttct	240
XX	Key	QY		
FT	Location/Qualifiers	241	agacaggtgttagatatacccaacttcttctgctctaaagcttactctgtctttgattgaa	300
FT	1..1404	Db		
FT	/*tag= a	241	agacaggtgttagatatacccaacttcttctgctctaaagcttactctgtctttgattgaa	300
FT	/product= "Phytase-1-thermo[8]-Q50T-K91A"	301	gctattcaaaagacgctactgtcttcaagggtaagtagcgttcttgaagacttacaac	360
XX	EP969089-AL.	Db		
XX	05-JAN-2000.	301	gctattcaaaagacgctactgtcttcaagggtaagtagcgttcttgaagacttacaac	360
XX	23-JUN-1999; 99EP-0111949.	361	tacacttgggtgctgacgactgtactccattcgttggaacacaaatggttaactctgt	420
XX	29-JUN-1998; 98EP-0111960.	Db		
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	361	tacacttgggtgctgacgactgtactccattcgttggaacacaaatggttaactctgt	420
XX	Brugger R, Lehmann M, Wyss M;	421	attlaagttctacagaagatacaagggcttggctagaaaagattgttccattcattagct	480
XX	WPI: 2000-099429/09.	Db		
XX	P-PSDB; AAY69568.	421	attlaagttctacagaagatacaagggcttggctagaaaagattgttccattcattagct	480
XX	New stabilized enzyme formulation, useful for feed compositions for monogastric animals -	481	tctggttctgacagagattattgtcttctgctgaaagtcttcaaggtttccaatctgt	540
XX	Example 5; Fig 19; 101pp; English.	Db		
XX	The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate	541	aagttggtgacccaggttctcaaccacacacacacacacacacacacacacacacac	600
XX	cc phosphorylase) and one or more stabilising agents including	Db		
XX	xyliol or ribitol; polyethylene glycols with a molecular weight of 600	601	ccagaaggtatccgggttac	660
XX	to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,	Db		
XX	glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.	661	tctgaattaggtgacgactgtgaagcttaacttctgctgaaagtcttcaaggtttccaatctgt	720
XX	The stabilised phytase formulation is used in a method for preparing a	Db		
XX	feed composition for monogastric animals (e.g., pigs, poultry) and	721	gctagattggaagctgacttgcaggtgttacttctgactgacgaagacgttgtttacttg	780
XX	provides a monogastric animal with its dietary requirements of	Db		
XX	phosphorus. Although a large amount of phosphate is present in animal	721	gctagattggaagctgacttgcaggtgttacttctgactgacgaagacgttgtttacttg	780
XX	feed in the form of phytate phosphorus, monogastric animals are unable	QY		
XX	to utilise this form of phosphate, resulting in the addition of extra	781	atgacatgtgtccattcgacactgtcgtagaacttctcagctactgactgaattgtctcca	840
XX	phosphate to the feed of such animals. Phytase enhances the nutritional	Db		
XX	value of plant material without the need for adding additional phosphate	781	atgacatgtgtccattcgacactgtcgtagaacttctcagctactgactgaattgtctcca	840
XX	to the feed. The level of phosphate pollution in the environment is	QY		
XX	reduced by adding phytase to animal feed, as the animal can make use of	841	tctgtgtttgttactccagcagaaatgattccaatacactacttctgcaaaagcttgggt	900
XX	the inorganic phosphate liberated from phytate phosphorus using the	Db		
XX	enzyme. The phytase formulation of the invention has an improved	841	tctgtgtttgttactccagcagaaatgattccaatacactacttctgcaaaagcttgggt	900
XX	thermostability and can therefore remain stable during long-term storage	QY		
XX	and can withstand feed processing methods such as extrusion, expansion	901	aagtactacggttaccggtgtgtgttaaccatttgggttcagctcaaggtgtgtttcgtct	960
XX	and pelleting. The present sequence represents DNA encoding a mutant	Db		
XX	phytase-1 consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a	901	aagtactacggttaccggtgtgtgttaaccatttgggttcagctcaaggtgtgtttcgtct	960
XX	temperature optimum and melting point 7 degrees Celsius higher than that	QY		
XX	of phytase-1 (AAY69558).	961	aacgaattgattgttagattgactcactctccagttcaagaccacacttctactaacac	1020
XX	Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;	Db		
XX	Query Match 100.0%; Score 1404; DB 21; Length 1404;	1021	actttggacttaaccagctacttctccattgaacgtacttctgttgaactgtacttctt	1080
XX	Best Local Similarity 100.0%; Pred. No. 0;	Db		
XX	Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	1021	actttggacttaaccagctacttctccattgaacgtacttctgttgaactgtacttctt	1080
QY	1 atgggcgtgttcgtcggtactgtccattgccaccttctggtttccacatccggtacc	QY		
Db		1081	cacgacacacatgatatctatttcttctggtttgttgaacggttaccacgcca	1140
		Db		
		1081	cacgacacacatgatatctatttcttctggtttgttgaacggttaccacgcca	1140
		1141	ttgtctactacttctgttgaattctattgaagaaactgacggttactctgtcttctt	1200

Db 781 atggacatgtccattcgacactgctcgtagaactcttgacgctactgaattgtctcca 840
 QY 841 ttctgtgctttgttctactcacacgaatggtatccaatcagactacttgcgaagcttgggt 900
 Db 841 ttctgtgctttgttctactcacacgaatggtatccaatcagactacttgcgaagcttgggt 900
 QY 901 aagtaactacggttacggtgctggtgaacccattgggtccagctcaaggtgttgggtttcgct 960
 Db 901 aagtaactacggttacggtgctggtgaacccattgggtccagctcaaggtgttgggtttcgct 960
 QY 961 aacgaattgattgtagattgactcaactctccagttccaggtccaaagcacttctactaaaccac 1020
 Db 961 aacgaattgattgtagattgactcaactctccagttccaggtccaaagcacttctactaaaccac 1020
 QY 1021 actttggactctaacccagctacttctccattgaaagcgtacttctggtttgtacagcttctct 1080
 Db 1021 actttggactctaacccagctacttctccattgaaagcgtacttctggtttgtacagcttctct 1080
 QY 1081 cagcaaacactatgatatactatttctctgctttggtttgtacacggttaccagcca 1140
 Db 1081 cagcaaacactatgatatactatttctctgctttggtttgtacacggttaccagcca 1140
 QY 1141 ttgttactactctctgttgatctattgaacaaactgacggttactctgttcttggact 1200
 Db 1141 ttgttactactctctgttgatctattgaacaaactgacggttactctgttcttggact 1200
 QY 1201 gttccattcgtctgtagctgttacgttgaaatgatgcaatgtcaagctgaaaggaacca 1260
 Db 1201 gttccattcgtctgtagctgttacgttgaaatgatgcaatgtcaagctgaaaggaacca 1260
 QY 1261 ttggttagagtttgggttaacacacagaggttgcctcattgacaggttgcgttgacaag 1320
 Db 1261 ttggttagagtttgggttaacacacagaggttgcctcattgacaggttgcgttgacaag 1320
 QY 1321 ttgggttagatgaagagacgactctgttgaaaggtttgtcttgcgttagatctgttgggt 1380
 Db 1321 ttgggttagatgaagagacgactctgttgaaaggtttgtcttgcgttagatctgttgggt 1380
 QY 1381 aactgggctgaattgttgcgttaa 1404
 Db 1381 aactgggctgaattgttgcgttaa 1404

RESULT 6

AAZ27423
 ID AAZ27423 standard; cDNA; 1426 BP.
 XX AC AAZ27423;
 XX DT 07-DEC-1999 (first entry)
 XX DE Ascomycete consensus phytase coding sequence.
 XX KW Phytase; variant; enzyme; phosphorus liberation; phytase substrate;
 KW phytate level reduction; animal manure; food preparation;
 KW soy processing; inositol manufacture; ss.
 XX OS Synthetic.
 XX PN W09949022-A1.
 XX PD 30-SEP-1999.
 XX PF 22-MAR-1999; 99WO-DK00153.
 XX PR 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX

PI Svendsen A;

XX WP1; 1999-580444/49.
 DR P-PSDB; AAY39906.
 XX

New variant phytase enzymes, used for liberating phosphorus from a
 phytase substrate, for reducing phytate levels in animal manure and in
 feed and food preparations -

XX Disclosure; Fig 9f-g; 14lpp; English.

XX This sequence encodes the consensus Ascomycete phytase sequence.
 CC The invention relates to variant phytase enzymes with specific amino acid
 CC substitutions for improved properties. The phytase variants can be used
 CC for liberating phosphorus from a phytase substrate. They can be used for
 CC reducing phytate levels in animal manure. They can be used in feed or
 CC food preparations. The phytase DNA can also be used to produce transgenic
 CC plants which can be used in feeds or foods. The phytase variants can also
 CC be used in soy processing and in the manufacture of inositol or
 CC derivatives. The phytase variants can have altered activities such as pH
 CC stability, temperature stability, pH profile, temperature profile,
 CC specific activity (in particular in relation to pH and temperature),
 CC substrate specificity, substrate cleavage pattern, substrate binding,
 CC position specificity, the velocity and level of release of phosphate
 CC from corn, reaction rate, phytase degradation rate and end level of
 CC released phosphate reached.

XX Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;

Query Match 96.9%; Score 1360.8; DB 20; Length 1426;
 Best Local Similarity 98.1%; Pred. No. 0;
 Matches 1377; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 atgggctgttcgtgctactgtccattgcccacttctgttcggttcacaccccggtacc 60
 Db 12 atgggctgttcgtgctactgtccattgcccacttctgttcggttcacaccccggtacc 71

QY 61 gcttgggttcctcgtggttaattctcactcttctgtgacactgttgacggttggtaccatgt 120
 Db 72 gcttgggttcctcgtggttaattctcactcttctgtgacactgttgacggttggtaccatgt 131

QY 121 ttcccagaaaatttctcacttctggtggttacctactctcactctctcttcttggagacgaa 180
 Db 132 ttcccagaaaatttctcacttctggtggttacctactctcactctctcttcttggagacgaa 191

QY 181 tctgtatttctccagacggttccagacgactgtagagttacttctgccaagtttctgt 240
 Db 192 tctgtatttctccagacggttccagacgactgtagagttacttctgccaagtttctgt 251

QY 241 agacacggtgctagataccacacttctctcgtctcgttaaggcttactctgtttgattgaa 300
 Db 252 agacacggtgctagataccacacttctctcgtctcgttaaggcttactctgtttgattgaa 311

QY 301 gctattcaaaaagacgctactgctttcgaaggttaagctacgcttcttctgaaagacttaaac 360
 Db 312 gctattcaaaaagacgctactgctttcgaaggttaagctacgcttcttctgaaagacttaaac 371

QY 361 tacactttgggtgctgacgactgactcattcgtgtaaaacacaaatgttgaactctggt 420
 Db 372 tacactttgggtgctgacgactgactcattcgtgtaaaacacaaatgttgaactctggt 431

QY 421 attagttctacagaagatacaaggctttgggtgtagaagattgttccattccattagagct 480
 Db 432 attagttctacagaagatacaaggctttgggtgtagaagattgttccattccattagagct 491

QY 481 tctggtttctgacagagttattgcttctgctgaaagttcattgaaggtttcccaactgct 540
 Db 492 tctggtttctgacagagttattgcttctgctgaaagttcattgaaggtttcccaactgct 551

QY 541 aagttgctgacccaggttctcaaccacacacaaagcttctccagttattacagtgatcatt 600
 Db 552 aagttgctgacccaggttctcaaccacacacaaagcttctccagttattacagtgatcatt 611

```

QY 601 ccagaagatccggtttacacacacatttgacacacggtactgtactgtcttcgaagac 660
    |||
    |||
    |||
Db 612 ccagaagatccggtttacacacacatttgacacacggtactgtactgtcttcgaagac 671
    |||
    |||
    |||
QY 661 tctgaattaggtgacacggtttgaagcttaacttcaactgtcttcttcgctccagctattaga 720
    |||
    |||
    |||
Db 672 tctgaattgggtgacacggtttgaagcttaacttcaactgtcttcttcgctccagctattaga 731
    |||
    |||
    |||
QY 721 gctagattggaagctgacttccaggtgtacttctgactgacgaagcgtttgttacttg 780
    |||
    |||
    |||
Db 732 gctagattggaagctgacttccaggtgtacttctgactgacgaagcgtttgttacttg 791
    |||
    |||
    |||
QY 781 atggacatgttccattcgacacgtctgctagaaacttctgacgtactgtaattgtctcca 840
    |||
    |||
    |||
Db 792 atggacatgttccattcgacacgtctgctagaaacttctgacgtactgtaattgtctcca 851
    |||
    |||
    |||
QY 841 ttctgtgttcttcaactcagacgaatggatccaatacgaactacttgcgaagcttgggt 900
    |||
    |||
    |||
Db 852 ttctgtgttcttcaactcagacgaatggatccaatacgaactacttgcgaatcttgggt 911
    |||
    |||
    |||
QY 901 aagtactacggttgaagctgtgctggttaacccattgggtccagctcaaggtgttgggttgcgt 960
    |||
    |||
    |||
Db 912 aagtactacggttgaagctgtgctggttaacccattgggtccagctcaaggtgttgggttgcgt 971
    |||
    |||
    |||
QY 961 aacgaattgattgctagattgactcactctccagttccagttcaagaccacacttctactaacac 1020
    |||
    |||
    |||
Db 972 aacgaattgattgctagattgactcactctccagttccagttcaagaccacacttctactaacac 1031
    |||
    |||
    |||
QY 1021 acttggactctaaacacgactacttccattggaacgctacttctgtaacgctgacttctct 1080
    |||
    |||
    |||
Db 1032 acttggactctaaacacgactacttccattggaacgctacttctgtaacgctgacttctct 1091
    |||
    |||
    |||
QY 1081 cagcaacacactatgatctatcttcttctgcttcttgggttctgtaacacggttaccagcca 1140
    |||
    |||
    |||
Db 1092 cagcaacacactatgatctatcttcttctgcttcttgggttctgtaacacggttactgctcca 1151
    |||
    |||
    |||
QY 1141 ttgtctactactcttctgtaactctattgaagaactgacggttactctgttcttggact 1200
    |||
    |||
    |||
Db 1152 ttgtctactactcttctgtaactctattgaagaactgacggttactctgttcttggact 1211
    |||
    |||
    |||
QY 1201 gttccattcgtctagagcttcaactgtgaatgtatgcaatgtcaagctgaaaggaaccca 1260
    |||
    |||
    |||
Db 1212 gttccattcgtctagagcttcaactgtgaatgtatgcaatgtcaagctgaaaggaaccca 1271
    |||
    |||
    |||
QY 1261 ttggttagagtttgggttaacacagaggttctccattgcacggttgggtgctgtgacacag 1320
    |||
    |||
    |||
Db 1272 ttggttagagtttgggttaacacagaggttctccattgcacggttgggtgctgtgacacag 1331
    |||
    |||
    |||
QY 1321 ttggttagagtttaagagagacgacttctggtgaaggttcttcttctgctagatctggtggt 1380
    |||
    |||
    |||
Db 1332 ttggttagagtttaagagagacgacttctggtgaaggttcttcttctgctagatctggtggt 1391
    |||
    |||
    |||
QY 1381 aactgggctgaatgttctcgttaa 1404
    |||
    |||
    |||
Db 1392 aactgggctgaatgttctcgttaa 1415
    |||
    |||
    |||
```

```

RESULT 7
AA23022
ID AAX23022 standard; DNA; 1426 BP.
XX
AC AAX23022;
XX
XX AAX23022;
XX
DT 11-JUN-1999 (first entry)
XX
DE Fungal phytase gene consensus DNA.
XX
KW phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
KW feed additive; variant; mutein; feed; pharmaceutical; ds.
XX
OS Fungi.
XX
```

```

FH Key Location/Qualifiers
FT CDS 12..1415
FT /*tag= a
FT /product= "phytase"
XX
PN EP897985-A2.
XX
PD 24-FEB-1999.
XX
PF 15-JUL-1998; 98EP-0113176.
XX
PR 24-JUL-1997; 97EP-0112688.
XX
PA (HOFF ) HOFFMANN LA ROCHE AG F.
XX
PI Lehmann M;
XX
DR WPI; 1999-134647/12.
DR P-PSDB; AAW93380.
XX
PT Preparation of a consensus protein, especially a phytase - using
    programs to compare evolutionary similarity of sequences
XX
XX Claim 8; Fig 2; 30pp; English.
XX
CC This invention describes a novel process for the preparation of a
    consensus protein. The specific example given in the specification is
    that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses
    phytase to valuable feed additives, with a fully defined amino acid
    sequence given in the specification, or variant or mutein. The method is
    useful for improving protein properties by altering their sequence. The
    consensus protein and mutein are useful in food, feed or pharmaceutical
    compositions. This sequence encodes the consensus phytase protein used
    in the method of the invention.
XX
SQ Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;
```

```

Query Match 96.9%; Score 1360.8; DB 20; Length 1426;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 atggcgctgttcgtctgactgtccattgccactgttcggttccacatccggtacc 60
Db 12 atggcgctgttcgtctgactgtccattgccactgttcggttccacatccggtacc 71
QY 61 gccctgggtcctcgtggttaattctcactctgtgacactgttgacggtggttaccatgt 120
Db 72 gccctgggtcctcgtggttaattctcactctgtgacactgttgacggtggttaccatgt 131
QY 121 ttcccagaataattctcactgtggtggttacactctccatactctcttcttggcagacgaa 180
Db 132 ttcccagaataattctcactgtggtggttacactctccatactctcttcttggcagacgaa 191
QY 181 tctgctatttctccagcgttccacagcactgtagacttacttcttcgaagtttctgtct 240
Db 192 tctgctatttctccagcgttccacagcactgtagacttacttcttcgaagtttctgtct 251
QY 241 agacacggtgctagatatacccaacttcttctgcttctaaagcttactctgtttgattgaa 300
Db 252 agacacggtgctagatatacccaacttcttctgcttctaaagcttactctgtttgattgaa 311
QY 301 gctattcaaaaagaacgctactgttctcaaggttaagtcagcttcttcttgaagacttacaac 360
Db 312 gctattcaaaaagaacgctactgttctcaaggttaagtcagcttcttcttgaagacttacaac 371
QY 361 tacactttgggtgctgacgacttgactcattcgttgaaacacaaatgggttaactctggt 420
Db 372 tacactttgggtgctgacgacttgactcattcgttgaaacacaaatgggttaactctggt 431
QY 421 ataatggtctacagaagatacaagcttgggtgctagaagattgttccattcattagacct 480
Db 432 ataatggtctacagaagatacaagcttgggtgctagaagattgttccattcattagacct 491
```


QY 301 gctattcaaaagacgctactgcttctcaagggtaagtaacgcttcttctgaagacttaaac 360
Db 312 gctattcaaaagacgctactgcttctcaagggtaagtaacgcttcttctgaagacttaaac 371
QY 361 tacactttgggtgctgacgacttactccttcggtgtaaaacacaaatggttaactctggt 420
Db 372 tacactttgggtgctgacgacttactccttcggtgtaaaacacaaatggttaactctggt 431
QY 421 attaatgtctacagaagatacaagcgtttggtctagaagattgttccattcattagagct 480
Db 432 attaatgtctacagaagatacaagcgtttggtctagaagattgttccattcattagagct 491
QY 481 tctgttctgacagaggttattgcttctgctgaaagttcattgaaggtttccaatctgct 540
Db 492 tctgttctgacagaggttattgcttctgctgaaagttcattgaaggtttccaatctgct 551
QY 541 aagtggctgacagcaggttctcaacacacacacacacacacacacacacacacacacac 600
Db 552 aagtggctgacagcaggttctcaacacacacacacacacacacacacacacacacacac 611
QY 601 ccagaaggaacccggttac 660
Db 612 ccagaaggaacccggttac 671
QY 661 tctgaattaggtgacagcgttgaagcacttactcctgcttctgctccagctattaga 720
Db 672 tctgaattaggtgacagcgttgaagcacttactcctgcttctgctccagctattaga 731
QY 721 gctagattggaagcgttactgacaggttacttgaactgacagcgttacttgaactg 780
Db 732 gctagattggaagcgttactgacaggttacttgaactgacagcgttacttgaactg 791
QY 781 atggacattgtccattcgcactgctgctgctgctgctgctgctgctgctgctgctgct 840
Db 792 atggacattgtccattcgcactgctgctgctgctgctgctgctgctgctgctgctgct 851
QY 841 tctgtgcttcttactcactcagcagcagcagcagcagcagcagcagcagcagcagcag 900
Db 852 tctgtgcttcttactcactcagcagcagcagcagcagcagcagcagcagcagcagcag 911
QY 901 aagtactcaggttactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 960
Db 912 aagtactcaggttactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 971
QY 961 aacgaattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1020
Db 972 aacgaattgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1031
QY 1021 actttgactcactcactcactcactcactcactcactcactcactcactcactcactcact 1080
Db 1032 actttgactcactcactcactcactcactcactcactcactcactcactcactcactcact 1091
QY 1081 cagcaacacactatgatatatatttcttctgctgctgctgctgctgctgctgctgctgctgct 1140
Db 1092 cagcaacacactatgatatatatttcttctgctgctgctgctgctgctgctgctgctgctgct 1151
QY 1141 ttgtctactctctgttgaactattgaaagcactgacggttactctctctctctgact 1200
Db 1152 ttgtctactctctgttgaactattgaaagcactgacggttactctctctctctgact 1211
QY 1201 gttccattcgtctgactgacttactgctgctgctgctgctgctgctgctgctgctgctgct 1260
Db 1212 gttccattcgtctgactgacttactgctgctgctgctgctgctgctgctgctgctgctgct 1271
QY 1261 ttggttagagtttgggttaacacagaggttgccttgcacaggttgccttgcacaggttgccttgcacag 1320
Db 1272 ttggttagagtttgggttaacacagaggttgccttgcacaggttgccttgcacaggttgccttgcacag 1331
QY 1321 ttggttagagtttgggttaacacagaggttgccttgcacaggttgccttgcacaggttgccttgcacag 1380
Db 1332 ttggttagagtttgggttaacacagaggttgccttgcacaggttgccttgcacaggttgccttgcacag 1391

QY 1381 aactgggctgaatgttgcgttaa 1404
Db 1392 aactgggctgaatgttgcgttaa 1415
RESULT 9
AAZ59637
ID AAZ59637 standard; DNA; 1426 bp.
XX AAZ59637;
AC AAZ59637;
XX 19-APR-2000 (first entry)
DT 19-APR-2000 (first entry)
XX DNA encoding phytase-1, a consensus phytase.
DE DNA encoding phytase-1, a consensus phytase.
XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
KW thermostable; animal feed; monogastric animal; phytate phosphorus;
KW phosphate availability; consensus; phytase-1; ds.
XX Aspergillus terreus 9A1.
OS Aspergillus terreus cbs16.46.
OS Aspergillus niger var. awamori.
OS Aspergillus niger T213.
OS Aspergillus fumigatus NRRL3135.
OS Aspergillus fumigatus ATCC13073.
OS Aspergillus fumigatus ATCC32722.
OS Aspergillus fumigatus ATCC58128.
OS Aspergillus fumigatus ATCC26906.
OS Aspergillus fumigatus ATCC32239.
OS Emericella nidulans.
OS Talaromyces thermophilus ATCC20186.
OS Myceliophthora thermophila.
OS Synthetic.
XX Key Location/Qualifiers
PH 12..1415
FT /*tag= a /product= "Phytase-1"
FT 12..89
FT /*tag= b
FT /*note= "Signal peptide from Aspergillus terreus cbs16.46"
FT mat_peptide 90..1412
FT /*tag= c /product= "Mature phytase-1"
XX EP969089-Al.
XX 05-JAN-2000.
XX 23-JUN-1999; 99EP-0111949.
XX 29-JUN-1998; 98EP-0111960.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Bruggen R, Lehmann M, Wyss M;
XX WPI; 2000-099429/09.
XX P-PSDB; AAY69558.
XX New stabilized enzyme formulation, useful for feed compositions for monogastric animals -
XX Example 3; Fig 14; 101pp; English.
XX The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate phosphohydrolase) and one or more stabilising agents including xylitol or ribitol; polyethylene glycols with a molecular weight of 600 to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic, glutaric and succinic acid; carboxymethylcellulose; and sodium alginate. The stabilised phytase formulation is used in a method for preparing a feed composition for monogastric animals (e.g., pigs, poultry) and

PN W0200043503-A1.
 XX 27-JUL-2000.
 PD 21-JAN-2000; 2000WO-DK00025.
 XX 21-JAN-1999; 99DK-0000092.
 PF 21-SEP-1999; 99DK-0001340.
 XX (NOVO) NOVO NORDISK AS.
 PA Lehmann M;
 XX WPI; 2000-491161/43.
 DR P-PSDB; AAB20527.
 XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -
 XX Claim 10; Fig 8a-c; 240pp; English.
 PS The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence encodes a phytase sequence from the present
 CC invention.
 XX Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;
 SQ

Query Match 94.5%; Score 1327.2; DB 21; Length 1404;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1356; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1 atggcggtgttcggtgactgtccattgacactgttgctggttccacatccggtacc 60
 Db 1 atggcggtgttcggtgactgtccattgacactgttgctggttccacatccggtacc 60

Qy 61 gccttgggtctcgtgtgtaattctcactctgtgacactgttgacactgttgacactgtg 120
 Db 61 gccttgggtctcgtgtgtaattctcactctgtgacactgttgacactgttgacactgtg 120

Qy 121 ttccagaaatttctcactgtggtggtgactctcactctcactctcactctcactctcact 180
 Db 121 ttccagaaatttctcactgtggtggtgactctcactctcactctcactctcactctcact 180

Qy 181 tctgctatttctcagacgttcacagcactgtagactgttagactgttagactgttagactgt 240
 Db 181 tctgctatttctcagacgttcacagcactgttagactgttagactgttagactgttagactgt 240

Qy 241 agacacggtgctagatgccaaactctctcgtgtcgtgtcgtgtcgtgtcgtgtcgtgtgaa 300
 Db 241 agacacggtgctagatgccaaactctctcgtgtcgtgtcgtgtcgtgtcgtgtgaa 300

Qy 301 gctattcaaaagacgactactgtttcaagggttaagtagcgtttcttgaagacttacac 360
 Db 301 gctattcaaaagacgactactgtttcaagggttaagtagcgtttcttgaagacttacac 360

Qy 361 tacacttgggtgctgacgactgtgactcattcgtgtggtgaaacaaatggttaactctggt 420
 Db 361 tacacttgggtgctgacgactgtgactcattcgtgtggtgaaacaaatggttaactctggt 420

Qy 421 attagttctacagaagatacagaagcttggctagagaagattgttccattcattagagct 480
 Db 421 attagttctacagaagatacagaagcttggctagagaagattgttccattcattagagct 480

Db 421 attagttctacagaagatacagaagcttggctagagaagattgttccattcattagagct 480
 Qy 481 tctggttctgacagagattgttctgctgaaagattcattgaaggtttccaatgtct 540
 Db 481 tctggttctgacagagattgttctgctgaaagattcattgaaggtttccaatgtct 540
 Qy 541 aagtggctgaccaggttctcaccacacacacacacacacacacacacacacacacacac 600
 Db 541 aagtggctgaccaggttctcaccacacacacacacacacacacacacacacacacacac 600
 Qy 601 ccagaagatccggttacaaac 660
 Db 601 ccagaagatccggttacaaac 660
 Qy 661 tctgaattggtgacgactgtgaagcttaacttcaactgcttctcactgcttctcactcact 720
 Db 661 tctgaattggtgacgactgtgaagcttaacttcaactgcttctcactgcttctcactcact 720
 Qy 721 gctagattggaagctgacttgcaggtgttaacttgaactgacgaagacgttcttacttg 780
 Db 721 gctagattggaagctgacttgcaggtgttaacttgaactgacgaagacgttcttacttg 780
 Qy 781 atggacatgttccattcgcactgtcgtagaacttctgacacttctgacactgactgacttct 840
 Db 781 atggacatgttccattcgcactgtcgtagaacttctgacacttctgacactgactgacttct 840
 Qy 841 tctgctgttctcactcagcgaatggtatccaatcagactacttgcacacgttgggt 900
 Db 841 tctgctgttctcactcagcgaatggtatccaatcagactacttgcacacgttgggt 900
 Qy 901 aagtactcgttgcgtgctgactgttaaccattgggtcagctcaggtgtggttctgct 960
 Db 901 aagtactcgttgcgtgctgactgttaaccattgggtcagctcaggtgtggttctgct 960
 Qy 961 aacgaattgattgctagattgactcactcctcagttcagacacacacttctactaacac 1020
 Db 961 aacgaattgattgctagattgactcactcctcagttcagacacacacttctactaacac 1020
 Qy 1021 accttggacttcaaccagctacttccattgaacgtacttctgacgtgacttctct 1080
 Db 1021 accttggacttcaaccagctacttccattgaacgtacttctgacgtgacttctct 1080
 Qy 1081 cagcaacacactatgatactcttcttctgcttctggttctgacacgttaccagcca 1140
 Db 1081 cagcaacacactatgatactcttcttctgcttctggttctgacacgttaccagcca 1140
 Qy 1141 ttgtctacttctgttgaatctattgaagaacactgacgttactctgcttcttggact 1200
 Db 1141 ttgtctacttctgttgaatctattgaagaacactgacgttactctgcttcttggact 1200
 Qy 1201 gttccattcgtcgtgactgacttgaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1260
 Db 1201 gttccattcgtcgtgactgacttgaatgaaatgaaatgaaatgaaatgaaatgaa 1260
 Qy 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttctgcttggact 1320
 Db 1261 ttggttagagtttgggttaacgacagagttgttccattgacggttctgcttggact 1320
 Qy 1321 ttggttagatgaaagacagacttctggttaaggttcttctgctagatctggttgg 1380
 Db 1321 ttggttagatgaaagacagacttctggttaaggttcttctgctagatctggttgg 1380
 Qy 1381 aactgggctgaatttctgcttaa 1404
 Db 1381 aactgggctgaatttctgcttaa 1404

RESULT 12
 AAZ59716
 ID AAZ59716 standard; DNA; 1404 BP.
 XX AC
 XX AC
 XX AAZ59716;
 XX

Job time: 5033 sec

Db 372 tacactttgggtgctgacgaacttgactccattcggtgaaacaaaaatggttaactctggt 431
Qy 421 attaatgtctacagaagatacaagcttttgctagaaagattttccattcattagagct 480
|||||
Db 432 attaatgtctacagaagatacaagcttttgctagaaagattttccattcattagagct 491
|||||
Qy 481 tctggtctgacagaggttatgtctctctgctgaaagttcattgaaggtttccaatctgct 540
|||||
Db 492 tctggtctgacagaggttatgtctctctgctgaaagttcattgaaggtttccaatctgct 551
|||||
Qy 541 aagtgtgctgacccaggttctcaacacacacagcttctccagttattaaacgtgatcatt 600
|||||
Db 552 aagtgtgctgacccaggttctcaacacacacagcttctccagttattaaacgttatatt 611
|||||
Qy 601 ccagaaggatccgggttacacacacactttggaccacggttactgttactgctttogaagac 660
|||||
Db 612 ccagaagggtgctggttacacacacactttggaccacggttactgttactgctttogaagaa 671
|||||
Qy 661 tctgaattaggtgacagaggttgaagcttaacttcaactgctttgttgcgtccagctattaga 720
|||||
Db 672 tctgaattgggtgacagaggttgaagcttaacttcaactgctttgttgcgtccacattaga 731
|||||
Qy 721 gctagattggaagctgacttgcaggtgttactttgactgacagagagctgtgttacttg 780
|||||
Db 732 gctagattggaagctcaacttgcaggtgttactttgactgacagagagctgtgttacttg 791
|||||
Qy 781 atggacatgtgtccattcgaacacttgcgtctagaaacttctgacgtactgaaattgtctcca 840
|||||
Db 792 atggacatgtgtccattcgaacacttgcgtctagaaacttctgacgtactgaaattgtctcca 851
|||||
Qy 841 tctctgtcttcttcaactcacagcaatggtatccaatcagactacttgcacagcttgggt 900
|||||
Db 852 tctctgtactgttcaactcacagcaatggtatccaatcagactacttgcacagcttgggt 911
|||||
Qy 901 aagtactacggttacggtgctgtgtaacccattgggtccagctcaaggtgtgtgttgcgt 960
|||||
Db 912 aagtactacggttacggtgctgtgtaacccattgggtccagctcaaggtgtgtgttgcgt 971
|||||
Qy 961 aacgaattgattgcttagattgactcaactctccagttcaagacacacacttctactaacccac 1020
|||||
Db 972 aacgaattgattgcttagattgactcaactctccagttcaagacacacacttctactaacccac 1031
|||||
Qy 1021 actttgactctaacccagctacttctccattgaaacgctactttgtacgctgacttctct 1080
|||||
Db 1032 actttgactctaacccagctacttctccattgaaacgctactttgtacgctgacttctct 1091
|||||
Qy 1081 caagacaacactatgatatactattttctctgctttgggtttgtacaacaggttaccagcca 1140
|||||
Db 1092 cagacaacactatggtttctattttctctgctttgggtttgtacaacggttaccagcca 1151
|||||
Qy 1141 ttgtctactacttctgttgaatctattgaagaaactgacggttactctctctcttggact 1200
|||||
Db 1152 ttgtctactacttctgttgaatctattgaagaaactgacggttactctctctcttggact 1211
|||||
Qy 1201 gttccattcgtctgtagagcttacgttgaatgatgcaatgtcaagctgaaaagggaacca 1260
|||||
Db 1212 gttccattcgtctgtagagcttacgttgaatgatgcaatgtgaaagctgaaaagggaacca 1271
|||||
Qy 1261 ttggttagaggttttgggttaacacagaggttgtccattgacaggttgtgtgtgacaag 1320
|||||
Db 1272 ttggttagaggttttgggttaacacagaggttgtccattgacaggttgtgtgtgacaag 1331
|||||
Qy 1321 ttggttagaggttaagagagagcgtctggtgaaaggtttgtcttctgctagatctggtggt 1380
|||||
Db 1332 ttggttagaggttaagagagagcgtctggtgaaaggtttgtcttctgctagatctggtggt 1391
|||||
Qy 1381 aactgggctgaatgtttcgtcttaa 1404
|||||
Db 1392 aactgggaagaatgtttcgtcttaa 1415
|||||

Patent No. 5221624

Sequence 28, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 75, Appl
Sequence 1, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 74, Appl
Sequence 74, Appl
Sequence 74, Appl
Sequence 4, Appl
Sequence 4, Appl

ALIGNMENTS

RESULT

US-09-121-425-3
; Sequence 3, Application US/09121425
; Patent No. 6153418
CENTRAL INTELLIGENCE

APPLICANT: Lehmann,

```

/ TITLE OF INVENTION: Consensus phytases
/
/ FILE REFERENCE: consensus phytases 13239
/
/ CURRENT APPLICATION NUMBER: US/09/121,425
/
/ CURRENT FILING DATE: 1998-07-23
/
/ EARLIER APPLICATION NUMBER: EPO 97112688.3
/
/ EARLIER FILING DATE: 1997-07-24
/
/ NUMBER OF SEQ. ID NOS.: 2
/

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1426
; TYPE: DNA

```

; ORGANISM: Artificial Sequence

Query Match	96.9%	Score 1360.8	DB 3	Length 1426
Best Local Similarity	98.1%	Pred. No. 0		
Matches 137	Conservative	0	Mismatches 27	Indels 0
Gaps				
1 atgggcgtgtcgtgctactgtccattgcacattgttcggttcacatcgcgtacc 60				
12 atgggcgtgtcgtgctactgtccattgcacattgttcggttcacatcgcgtacc 71				
61 gcccttgggttcctcgtgtaatttcacattctgtgcacactgttgacggtgggtaccaatgt 120				
72 gcccttgggttcctcgtgtaatttcacattctgtgcacactgttgacggtgggtaccaatgt 131				
121 ttccagaaattctcacitgtgggtaccactactctcatctctctttggcagacgaa 180				
132 ttccagaaattctcacitgtgggtccaactactctcatctctctttgggaagacgaa 191				
181 tctgctatttctccagaactgtccagacactgtagagtctcttcgttccaagtttgcct 240				
192 tctgctatttctccagaactgtccagacactgtagagtctcttcgttccaagtttgcct 251				
241 agacacggtcctagataccaccaactcttctcgtcttaaggcttactcgtcttgattgaa 300				
252 agacacggtcctagataccaccaactcttcttaagtctaaggcttactcgtcttgattgaa 311				
301 gctattccaaagaacgtactcgtttccaagggtaagtacgctcttcttgaagacttacaac 360				

Db 312 gctattcaaaagacgctactgcttttcaagggttaagtacgcttttcttgaagacttaaac 371
Qy 361 tacactttgggtgctgacgacttgaactccattcggtgaaacacaaatgggttaactctgg 420
Db 372 tacactttgggtgctgacgacttgaactccattcggtgaaacacaaatgggttaactctgg 431
Qy 421 attaatgtctacagaagatacaagcttggctagaaagattgttccattcattagact 480
Db 432 attaatgtctacagaagatacaagcttggctagaaagattgttccattcattagact 491
Qy 481 tctggttctgacagagttattgcttctgctgaaagttcattacaaagtttcccaactctg 540
Db 492 tctggttctgacagagttattgcttctgctgaaagttcattacaaagtttcccaactctg 551
Qy 541 aagttggctgacagagtttctcaaccacacacaaagcttctccagttattaaacgtgatac 600
Db 552 aagttggctgacagagtttctcaaccacacacaaagcttctccagttattaaacgtgatac 611
Qy 601 ccagaagatccggttcaac 660
Db 612 ccagaagatccggttcaac 671
Qy 661 tctgaattagtgacgacgttgaagcttaacttcaacttctgcttctgctccagctattaga 720
Db 672 tctgaattgggtgacgacgttgaagcttaacttcaacttctgcttctgctccagctattaga 731
Qy 721 gctagattggaagctgacttgcaggtgttacttctgactgacgacgacgttgtttacttg 780
Db 732 gctagattggaagctgacttgcaggtgttacttctgactgacgacgacgttgtttacttg 791
Qy 781 atggacatgttccattcgacacacttgcctagaaacttgcctgacacacacacacacacac 840
Db 792 atggacatgttccattcgacacacttgcctagaaacttgcctgacacacacacacacacac 851
Qy 841 tctgtgtttgttcaactac 900
Db 852 tctgtgtttgttcaactac 911
Qy 901 aagttactcgttgaagctgctggttaacccattgggtccagctcaggtgttgggttggct 960
Db 912 aagttactcgttgaagctgctggttaacccattgggtccagctcaggtgttgggttggct 971
Qy 961 aacgaattgattgtagattgactcacttccagttccagttccagttccagttccagttccag 1020
Db 972 aacgaattgattgtagattgactcacttccagttccagttccagttccagttccagttccag 1031
Qy 1021 acttggacttaacccagctacttccagttccagttccagttccagttccagttccagttccag 1080
Db 1032 acttggacttaacccagctacttccagttccagttccagttccagttccagttccagttccag 1091
Qy 1081 cagcaac 1140
Db 1092 cagcaac 1151
Qy 1141 ttgtctactacttctgttgaattctattgaagaactgacggttactctgtcttcttggact 1200
Db 1152 ttgtctactacttctgttgaattctattgaagaactgacggttactctgtcttcttggact 1211
Qy 1201 gttccattcgtctgtagacttactgtgaattgactgacacacacacacacacacacacacac 1260
Db 1212 gttccattcgtctgtagacttactgtgaattgactgacacacacacacacacacacacacac 1271
Qy 1261 ttggttagagtttgggttaac 1320
Db 1272 ttggttagagtttgggttaac 1331
Qy 1321 ttgggttagatgtaagagacgacttctgttgaaggttcttcttctgctagactctgttgggt 1380
Db 1332 ttgggttagatgtaagagacgacttctgttgaaggttcttcttctgctagactctgttgggt 1391
Qy 1381 aactgggtgaaatgttctgcttaa 1404
Db 1392 aactgggtgaaatgttctgcttaa 1415

RESULT 2
US-08-151-574-33
; Sequence 33, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F. M. Van Gorkom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G. M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; US-08-151-574-33

Query Match 45.2%; Score 635; DB 1; Length 1404;
Best Local Similarity 65.8%; Pred. No. 5.8e-190;
Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;
Qy 1 atggggcggttcgctgctgactgctccattggccaccttggcttccggttccacacatccggtacc 60
Db 1 ATGGGCGGTCTCTGCTGTCTACTTCTCTTGTATCTCTCTGTCTGAGTCACCTCCGGACTG 60
Qy 61 gccttgggttcctcggtgtaattctcactcttctgacactgttgacggtgttaccatgt 120
Db 61 CGAGTCCCGGCTCGAGANAATCAATCCAGITGCGATACGGTCGATCAGGGGTATCAATGC 120
Qy 121 ttccacagaattctcacttgggtgtacacttctccatacttctcttcttggcagacgaa 180
Db 121 TTCTCCGAGACTTCGCATCTTTGGGGTCAATACGACCGTCTTCTCTCTCTGGAACGAA 180

QY 181 tctgtatttctccagacgcttccagacgactgtagatttcttcttccggttcaagtttgtct 240
Db 181 TCGGTCATCTCCCTAGGTGCCCGCGGATGACAGATCACTTTCGCTCAGGTCCTCC 240
QY 241 agacacggtgctagatacccaacttcttctgctcctaaagcttactcttctgattgaa 300
Db 241 GTCATGAGCGCGGTATCCGACCGACTTCAAGGCGAAGAAATACTCCGCTCATGTAG 300
QY 301 gctattcaaaagacgcttactgttttcaaggttaagtacgcttcttcttgaagacttaaac 360
Db 301 GAGATCCAGCAGACGCGACCACTTTGACGGAAATATGCTTTCCTGGAAGACATAAC 360
QY 361 tacactttggtgctgacgacttacttccattctcgttgaaaccaaattggttaactctggt 420
Db 361 TACAGCTTGGGTGAGATGACCTGACTCCCTTCGGAACAGGAGCTAGTCAACTCCGGC 420
QY 421 attaatctacagaagaagcttctgctcctcctgctcctcctcctcctcctcctcctcct 480
Db 421 ATCAAGTCTACAGCGGTAGGAATCGCTCACAAGGAACATCTCTCATTCATCCGATCC 480
QY 481 tctgtttctacagagatttattgtcttctgctgaaagtctcattgaaggtttccaaatcgt 540
Db 481 TCTGGCTCCAGCGCGTATCGCTCCGCGCAAGAAATTCATCCAGGGCTTCCAGACACC 540
QY 541 aagtgtctacccaggttctcacaacacacacacacacacacacacacacacacacacac 600
Db 541 AAGCTGAAGATCTCTGTCGCCAGCCCGGCAATCTCGCCCAAGATCGACGTGGTCAAT 600
QY 601 ccagaagattccggttacaacaacactttgaccacggttacttctgactcttccgaagac 660
Db 601 TCCGAGCCAGCTCATCCAACACACTCTGACCCAGGACCTGCTCTTCGAAGAC 660
QY 661 tctgaattaggtgacagcttgaagtaacttcaactgttcttctcctcctcctcctcctc 720
Db 661 AGCGAATGGCCGATACCGTCGAAGCAATTCACCGCCACGCTCTGCTCCCTCCATCTCT 720
QY 721 gctagattgaagctacacttgcaggttacttcttctgactgaagacgcttcttacttg 780
Db 721 CAAGCTGGAGAACGACTGTCGGGTGTGACTCTCACAGACACAGAGTACCTACCTC 780
QY 781 atggacattgtctcattccacacgctcgcgtagaaactctgacgactgaattgtctcca 840
Db 781 ATGGACATGTGCTCTTGACACCACTCTCCACCAGCACCGTCTGACACCAAGCTGCCCC 840
QY 841 tctgtgcttctcactcagcagcaatggtatccaatcagcacttctgcaagcttgggt 900
Db 841 TTCTGTGACTGTTCACCATGCAATGGATCACTACCTACCTACCTCAGCTCTTGA 900
QY 901 aagctactcaggttactggttgaacccattgggtcagctcagaggttgggttgcgt 960
Db 901 AAGTATTACGGCCATGGTCAGGTAAACCGCTCGGCCCGACCCAGGGGCTCGGCTACGCT 960
QY 961 aacgaattgattgcttagtactcactcctcagttcagttcagacacacttctactaacac 1020
Db 961 AACGAGCTCATCGCGCTGTGACCCACTCGCTGTCCACGATCACACCAAGTTCACAC 1020
QY 1021 acttggactctaacccagctacttccacttgaagcgttacttgaagcgtgacttctct 1080
Db 1021 ACTTTGGACTCGAGCCGCTACCTTTCCGCTCACTCTACTCTCTACGCGGACTTTTCG 1080
QY 1081 cagcaacaactatgatctcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
Db 1081 CATGACACGGCATCATCTCCATCTCTTTGCTTTAGGTCTGTACACGGCACTAAGCG 1140
QY 1141 ttgttactacttcttcttgaattcttattgaagaactgacggttactctgcttcttggact 1200
Db 1141 CTATCTACACGACCGTGGAGAAATATCACCCAGACAGATGATCTCTCTGCTGTGACG 1200
QY 1201 gttcattctcgtctagagcttactgaaatgctgaatcgaatcgaatcgaatcgaatcga 1260
Db 1201 GTTCGCTTGTCTCGCTTTGTAGCTCGAGATGATGCACTGCTAGGGGGGAGCAGGCG 1260
QY 1261 ttggttagagtttggtttaacgacagagatttcttccattgacggttctgctgttgacaag 1320

Db 1261 CTGGTCCGCTGTCTGGTTAATGATCGCTGTTCGGCTCATGGGTGTCCGCTTGATGCT 1320
QY 1321 ttgggttagatgaagagacgacgacttctggtgaaggttcttcttctgactagacttgggt 1380
Db 1321 TTGGGAGATGTACCCGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGT 1380
QY 1381 aactgggtgaattgttcttcttca 1403
Db 1381 GATTGGCGGAGTGTTCCTTA 1403

RESULT 3

US-08-146-424-19
; Sequence 19, Application US/08146424
; Patent No. 5593963
; GENERAL INFORMATION:
; APPLICANT: VAN OIJEN, ALBERT J. J.
; APPLICANT: RIETVELD, KRJUN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/146,424
; APPLICATION NUMBER: 33,407
; FILING DATE: 02-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KENNEDY, BILL
; REGISTRATION NUMBER: 33,407
; REFERENCE/DOCKET NUMBER: 44615-20011.24
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1401
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 70
; US-08-146-424-19

Query Match 45.2%; Score 635; DB 1; Length 1404;
Best Local Similarity 65.8%; Pred. No. 5.8e-190;
Matches 923; Conservative 0; Mismatches 480; Indels 0; Gaps 0;
QY 1 atggcggtgtctctgctgctactgtccattgcaactgttctggttccacatccggtacc 60
Db 1 ATGGCGGTCTCTGCTGTCTACTCTCTTTGTATCTCTGTGAGATCACCTCCGGACTG 60
QY 61 gcttgggtctctcgtggttaattctcactcttctgacactgttgacggtgttaccatgt 120

Db 61 GCAGTCCCGGCTCGAGAAATCAATCAGTTGGATACGTCGATCAGGGGTATCAATGC 120
Qy 121 ttccagaaatttctacttgggttacactctccatacttctcttggcagacgaa 180
Db 121 TTCGCCGAGACTTCGCATCTTGGGTCATATACGACCGTCTTCTCTGCGAAGAA 180
Qy 181 tctgtatttccagagcttccagagacttagagttacttcttgcgttaagtttgcct 240
Db 181 TCAGTCATCTCCCTGAGTGGCGCGGATGAGAGTCACTTTCGCTCAGGTCCTCTCC 240
Qy 241 agacaggtgtagatatacccaacttcttgcgtctaaggtcttacttctgttgaattgaa 300
Db 241 CGTCATGGAGCGCGGTATCCGACCGATCCCAAGGGCAAGAAATACTCCGCTCTCATGAG 300
Qy 301 gctattcaaaagacgtactcttccaaaggttaagtaagcttcttgaagacttaaac 360
Db 301 GAGATCCAGCAGAACGCGACCACTTTGACGGGAAATATGCTTCCTGAAGACATACAAC 360
Qy 361 tacactttgggtgctgacgacttactccattcgttggtgaaacccaaatggttaactctggt 420
Db 361 TACAGCTTGGGTGAGATGACTTACTTCCCTTCGAGAACAGGAGCTAGTCAACTCCGGC 420
Qy 421 attaatgtctacagagatacaaggttggctagaagatttccattcatttagagct 480
Db 421 ATCAAGTTTCTACACGCGTACGAATCGCTCAAGGACACATCGTTCTTCAATCCGATCC 480
Qy 481 tctgttctgacagagttatttcttctgctgaaagttcattgaagtttccaaatctgct 540
Db 481 TCTGCTCCAGCGGCTGATCGCTCCGCGCAAGAAATTCATCGAGGGCTTCCAGAGCACC 540
Qy 541 aagttggctgacccaggttctcaaccacacacagcttctccagttatttaacgtgactatt 600
Db 541 AAGCTGAAGGATCTCTGCTGCGCAGCGCGCAATCGTGCGCCCAAGATCGAGCTGGTCATT 600
Qy 601 ccagagagatccggttcaacaaacttggaccacggttacttactgcttcttgaagac 660
Db 601 TCCGAGGCCAGCTATCAACAACTCTCCAGCCAGCACCCTGCACCTGTCTTCCGAGAC 660
Qy 661 tctgaattagtgacagcgttgaagtaacttcaactgcttcttggctccagctattaga 720
Db 661 AGCGAATTTGGCGATACCTGCGAGCCAAATTTACCGCCAGCTTCTGCTCCCTCCATCGT 720
Qy 721 gctagattgaagctgaacttccaggttacttctgactgacgagacgttcttacttg 780
Db 721 CAACGCTCTGAGAGCAGCTGCTCGGCTGTGACTCTCAGACACAGAACTGACCTACCTC 780
Qy 781 atggacatgttccattcgacactgctgtagaacttctgacgtactgaattgtctcca 840
Db 781 ATGGACATGTCTCTCGACACCATCTCCACAGCACCGTCGACACCAAGCTGTCCGCC 840
Qy 841 tctgtcttggcttactcaacagcaaatggatcccaatacagactacttgcgaagcttgggt 900
Db 841 TTCCTGTGACTCTTCAACCATACGAATGGATCAACTACGACTACTCTCAGTCCCTTGA 900
Qy 901 aagtaactcaggttactgctgtgtaaccattgggtccagctcaaggttgggttgcct 960
Db 901 AAGTATTACGGCCATGGTCAGGTAACCGCTCGGCCCGCCAGCCGCTCGGTACGCT 960
Qy 961 aacgaattgattgttagtactcacttccacttccagttcagacacacacttctactaacac 1020
Db 961 AAGGAGCTCATCGCGCTGTGACCCACTCGCGCTGCTCCAGATGACACCAAGTTCACACAC 1020
Qy 1021 actttgacttaacccagctacttccattccattgaagcttacttctgactgacttctct 1080
Db 1021 ACTTTGGATCGAGCCGCTACTTCTTCGCTCACTTACTCTCTACGCGACTTTTCG 1080
Qy 1081 cacgaacacactatgatctatttcttctgcttgggtttgttacaacggttaccagcca 1140
Db 1081 CATGACAACGGATCATCTCCATCTCTTGTCTTAGGCTCTTACAAACGGCACTAAGCG 1140
Qy 1141 ttgtctactacttctgttgatctattgaagaactgacggttacttctgtcttggact 1200

Db 1141 CTATCTACCACCGCTGAGAAATATCACCAGACAGATGATTCTCGTCTGCTTGAAGC 1200
Qy 1201 gttccattcgtcgttagagcttaagttgaaatgatgaaatgtaagctgaaagaaacca 1260
Db 1201 GTTCGCTTGTCTCGGTTTGTACGTCGAGATGATGCAAGTGTACGCGGAGACGAGCGC 1260
Qy 1261 ttggttagattttggtttaacgacagagtttccattgcaaggttctgctgttgaacaag 1320
Db 1261 CTGGTCCGCTGCTTGGTTAATGATCGCGTGTTCGCGCTGATGGGTGTCGCGTTGATGCT 1320
Qy 1321 ttgggttagatgtaagagagacgacttctgtaaggtttgttcttctcgttagatctcgttgg 1380
Db 1321 TTGGGAGATGTAACCGGATAGCTTCTGAGGGGTGAGCTTGTCTAGATCTCTGGGGGT 1380
Qy 1381 aactgggtgaaatgttctcgtta 1403
Db 1381 GATTGGCGGAGTGTCTTGTCTTA 1403

RESULT 4
US-08-693-709-1
; Sequence 1, Application US/08693709
; Patent No. 5770413
; GENERAL INFORMATION:
; APPLICANT: VAN OIJEN, ALBERT J.J.
; APPLICANT: RIETVELD, KRIJN
; APPLICANT: HOEKEMA, ANDREAS
; APPLICANT: PEN, JAN
; APPLICANT: SIJMONS, PETER C.
; APPLICANT: VERWOERD, TEUNIS C.
; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,709
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146,424
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20011.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1401
; OTHER INFORMATION:
; NAME/KEY: mat.peptide
; LOCATION: 70...1401
; OTHER INFORMATION:

0.

; PRIOR APPLICATION DATA:

```

> APPLICATION NUMBER: 07/688, 578
> FILING DATE: 24-MAY-1991
> ATTORNEY/AGENT INFORMATION:
> NAME: Murashige, Kate H.
> REGISTRATION NUMBER: 29,959
> REFERENCE/DOCKET NUMBER: 24615-20026.00
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 415-327-7250
> INFORMATION FOR SEQ ID NO: 31:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 6756 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: double
> TOPOLOGY: linear
> MOLECULE TYPE: DNA (genomic)
> HYPOTHETICAL: NO
> ANTI-SENSE: NO
> ORIGINAL SOURCE:
> ORGANISM: Aspergillus ficuum (Aspergillus niger)
> STRAIN: NRL 3135
> IMMEDIATE SOURCE:
> LIBRARY: lambda AF
> CLONE: pAF2-3, pAF2-6, pAF2-7
> FEATURE:
> NAME/KEY: exon
> LOCATION: 210..253
> FEATURE:
> NAME/KEY: intron
> LOCATION: 254..355
> FEATURE:
> NAME/KEY: exon
> LOCATION: 356..1715
> FEATURE:
> NAME/KEY: CDS
> LOCATION: join(210..253, 356..1715)
> OTHER INFORMATION: /codon_start= 210
> OTHER INFORMATION: /product= "Phytase"
> FEATURE:
> NAME/KEY: sig_peptide
> LOCATION: 210..380
> FEATURE:
> NAME/KEY: mat_peptide
> LOCATION: 381..1712
> IDENTIFICATION METHOD: experimental
> OTHER INFORMATION: /function= "inositol phosphate"
> OTHER INFORMATION: phosphate
> OTHER INFORMATION: /product= "Phytase"
> OTHER INFORMATION: /evidence= EXPERIMENTAL
> LS-08-151-574-31

```

Query Match	44.9%;	Score 631;	DB 1;	Length 6756;
Best Local Similarity	65.9%;	Pred. No. 2.5e-108;		
Matches 916;	Conservative 0;	Mismatches 475;	Indels 0;	Gaps
QY	13	gtcgtgtactgtccattgcacacttgttcggtttccacatccggtacgcgttgggtcct	72	
Db	324	GTGGGACTTACGTGCGTGACTTATCTGTGCAGAGTCACTCTCGGACATGGCAGTCCCGCGC	383	
QY	73	cgtggtaattctcaactcttgtgacactgttgacggtgggtaccgaattgtccacagaatt	132	
Db	384	TCGAGAAATCAATCCAGTTCCGATACGGTCATGATCAGGGGTATCAATGCTTCTCGAGACT	443	
QY	133	tctcaactgtgggtacctactctctcaactctctctcttcttggcagacgaatctgtattct	192	
Db	444	TGGCATCTTTGGGGTCAATACGCACCGTTTCTCTCTGTGCAAAAGAAATCGGTCACTCC	503	
QY	193	ccagacgttccacagcactgtagagttacttccgttcaagtttgtctagaacagggtct	252	
Db	504	CCTGAGGTGCCCGCGATCCAGAGTCACTTTCGCTCAGGTCCTTCTCCCGTCATGAGGCG	563	
QY	253	agatacccaactcttcttgctaaagcttactctgtcttggatgaagctattccaaga	312	

Db	564	CGGTATTCGACCGACTCCAAAGGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG	621
QY	313	aacgctactgctttcaagggtaagctacgcttcttgaagacttaacactacactttgggt	372
Db	624	AACGGCACCACTTTGACGGGAAATATGCGCTTCTGAAGACATACAACTTACAGCTTGGGT	683
QY	373	gctgacgacttgactccattcgttgaaacacaaatggttaactctgttattaaagtctac	432
Db	684	GCAGATGACCTGACTCCCTTCGGAGAACAGGAGTAGTCAACTCCGGCATCAAGTTCTAC	743
QY	433	agaagatacaaggtttggctagaaagattgtccattcattagagctctcgtgtctgtac	492
Db	744	CAGCGGTACGAATCGCTCAACAGGAACATGTTCCATTATCCGATCTCTGGCTCCAGC	803
QY	493	agagttattgcttctgtgaaagtctaatgaaggtttccaatctgctaagttgctgac	552
Db	804	CGCGTGTACGCTTCGGCACAAGAAATTCATCGAGGGCTTCAGAGAGCACCAAGCTGAAGAT	863
QY	553	ccaggttctaacacacaccaggcttctccagttattaaactgtgatcattccagaagatcc	612
Db	864	CCTCGTCCACAGCCCGGCCAATCGTCGCCCAAGATCGACGTGTCATTTCGAGGCCAGC	923
QY	613	ggttacacaacactttggacacaggtacttgtactgtcttogaagactctgaattaggt	672
Db	924	TCATCCAAACAACATTCGACCCAGGACCACTGCACCTGTCTTCCAGACAGGGAATTTGCC	983
QY	673	gacgactgtaagtaacttcaactctgcttctgtctgcacgctattagagctagattggaa	732
Db	984	GATACGTCGAAGCAATTTACCGGCCAGTTCGTGCTCCCTCCATTTCGTCAACGCTTGAG	1043
QY	733	gctgactgccaaggtgttactttgactgacgaagcgttgtttacttgtatggacatggt	792
Db	1044	AACGACCTGTCCGCTGTGACTCTCAGACACAGAGAAGTGACCTACCTCATGGACATGTC	1103
QY	793	ccattcaacactctgctagaactctgacgtactgtaagtgtctccattctcgtgtttg	852
Db	1104	TCTTTCGACACCAATCTCCACGACCGCTGACACCAAGCTGTCCGCCCTTCTGTGACCTG	1163
QY	853	ttcactcacgacgaatggattccaaactacgactacttgcaaaagcttgggttaagtactacggt	912
Db	1164	TTACCCCATGACGAATGGATCACTACGACTACCTTCCAGTCTCTTGAANAAGATTACGGC	1223
QY	913	tacgctgctggtaaccacttgggtccagctcaaggttgtttctgcgtcaacgaattgatt	972
Db	1224	CATGGTCAGGTAAACCGGTTCGGCCCGACCCAGGGCGTCGGCTACGCTAACGAGCTCATC	1283
QY	973	gctagattgactcaactctccagttcaagacacactctactaacacacactttggactct	1032
Db	1284	GCCGCTGTGACCCACATCGCGTGTCCACGATGACACCACTTCCACCAACACTTTGGACTCG	1343
QY	1033	aaccagctactttccattgaacgctactttgtacgctgacttctcgcagacaacact	1092
Db	1344	AGCCGGCTACCTTTCCGCTCAACTCTACTCTACGCGGACTTTTCGATGACACGGC	1403
QY	1093	atgatatctatttcttcgctttgggtttgtacaacggtaccaagccattgtctactact	1152
Db	1404	ATCATCTCCATTCTCTTTCCTTAGTGCTGTACAAGGCATAAGCCGCTATCTACCAAG	1463
QY	1153	tctgtgaactctattgaagaaactgaacgttactctgcttcttgactgttccacttgcct	1212
Db	1464	ACCGTGGAGNATATCACCAGACACATGGATTCTCGTCTGCTTGGACGGTTCGGTTTGCT	1523
QY	1213	gctagagcttaagttgaatgatgaatgtcaagctgaaggaaggaaccattgggttagagt	1272
Db	1524	TCGCGTTTGTACGTCAGATGATGACGTGTACGGCGGACAGGAGCCCTGGTCCCGTGTC	1583
QY	1273	ttggttaacgacagagttgttccattgcacggttgtcgtttgacaagttgggttagatgt	1332
Db	1584	TTGGTTAATGATCGCGTTCGCCGTGCATGGGTCTCGGCTTGATGCTTTGGGGAGATGT	1643
QY	1333	aagagagacgacttctgtgaaggtttgtcttctgcctagactcgtgtgaactcgggtgaa	1392
Db	1644	ACCCGGATAGCTTTGTAGGGGGTTGAGCTTCTAGATCTGGGGGTGATTGGCGGAG	1703

```

; NAME/KEY: sig_peptide
; LOCATION: 210..380
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 381..1712
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "inositol phosphate
; OTHER INFORMATION: /product= "Phytase"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; US-08-419-448-31

Query Match 44.9%; Score 631; DB 2; Length 6756;
Best Local Similarity 65.9%; Pred. No. 2.5e-188;
Matches 916; Conservative 0; Mismatches 475; Indels 0; Gaps 0;

Qy 13 gctgtactgtccattgccacattgttgcgttccacatccggtaccgcttgggtcct 72
Db 324 GTGGGACTACTGATCGCTGACTATCTGTGACAGAGTCACTCCGAGTCCGCCGCC 383

Qy 73 cgtgtaattctcactcttgaacactgttgacggtgttaccacatgttccacgaatt 132
Db 384 TCGAGAAATCAATCCAGTTGCGATACGGTGCATCAGGGGTATCAATGCTTCTCCGAGACT 443

Qy 133 tctcacttgtgggtacacttctccatcttctcttcttggcagacgaattctgtattct 192
Db 444 TCGCATCTTTGGGGTCAATACGCCCGGTTCTCTTGGCAACGAATCGGTCTATCTCC 503

Qy 193 ccagacgttccagacagactgttagagttacttccgttccaaagttttcttagacacgtgtct 252
Db 504 CTTGAGGTGCCCGCGATGTCAGAGTCACTTTTCGTCTAGGTCTCTCCCGTCTGAGGCG 563

Qy 253 agataccacacttcttctgcctgaaggttactctgttctgttgaagcatttcaaaag 312
Db 564 CGGTATCCGACCGACTCCCAAGGGCAGAAATATCTCGCTCTCATTTGAGGAGATCCAGCAG 623

Qy 313 aacgtactgtttcaagggtgaagtacgcttcttgaagacttacaactacacttgggt 372
Db 624 AACGCGACCCACTTTGACGGAAAAATATGCTTCTCTGAACACATACAACTACAGCTTGGGT 683

Qy 373 gctgacgttgaactccattcggtagaaacccaaagttaactctgttatttaagttctac 432
Db 684 GCAGATGACCTGACTCTCCCTTCGAGAACAGAGGAGTGTAGTCAACTCCGGCATCAAGTTCTAC 743

Qy 433 agaagatacgaagcttggctagaagaattgttccattcattagagcttctgttctgac 492
Db 744 CAGCGGTACGAATCGCTCACAAGGAACACTGTTCCATTTCATCCGATCTCTGSCITCCAGC 803

Qy 493 agagtattgtcttctgtaaaagtccattgaagggtttccaaatctgtcgaagtgtgac 552
Db 804 CGCGTGATCGCTCCGGCAAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGGAT 863

Qy 553 ccaggttctcaaccacacacacacacacacacacacacacacacacacacacacacac 612
Db 864 CCTCGTCCCGACCGCCGCCAATCGTCCGCCAAGATCGACGTGTCTATTTCCGAGGCCAGC 923

Qy 613 ggttacaacacacacttggaccacggttactgttactgttcttgaagacactctgaattaggt 672
Db 924 TCATCCAAACAACACTCTCGACCCCGGACCTGTCACACTGTCTCTCGAAGACAGCGAATGGCC 983

Qy 673 gacgaogtgaagacttaactcactgttctgttgcgtccagctatttagagctagattgaa 732
Db 984 GATACCGTCGAAGCAATTTTACCGCCACCTGCTGCTCCCTCCCTTCCATTCGTCAACGTCGTGGAG 1043

Qy 733 gctgacttgcaggtgttacttgaactgacgaagacggttgttacttgaatgacatgtgt 792
Db 1044 AACGACCTGTCCGGTGTGACTCTTCACAGACACAGAGTGAACCTACCTCATGGACATGTC 1103

Qy 793 ccattcgacactgtcgttagaacttctgacgtactgaattgtctcattctgtgttctg 852
Db 1104 TCCTTCGACACCATCTCCACGACCGTCCGACACCAAGGTGTCTCCCTCTCTGTGACCTG 1163

```

```

Qy 1393 tgttgcctta 1403
Db 1704 TGTTCGCTTA 1714

RESULT 7
US-08-419-448-31
; Sequence 31, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorkom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paillon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Seltin
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murshige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; IMMEDIATE SOURCE:
; LIBRARY: lambda AF
; CLONE: pAF2-3, pAF2-6, pAF2-7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 210..253
; FEATURE:
; NAME/KEY: intron
; LOCATION: 254..355
; FEATURE:
; NAME/KEY: exon
; LOCATION: 356..1715
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(210..253, 356..1715)
; OTHER INFORMATION: /codon_start= 210
; OTHER INFORMATION: /product= "Phytase"
; FEATURE:

```

```

Qy 853 ttcaactcagcagaatgatccaatacagactacttgcaaaagcttgggtaagtaactacgggt 912
Db 1164 TTCACCCATGAGATGATCAATACACGATACCTCCAGTCTTGAAGAAAGTATTACGGC 1223
Qy 913 tacggtgctgtaacccattgggtccagctcaaggtgtgttctcgttaacgaattgatt 972
Db 1224 CATGCTGAGGTAACCCGCTCGGCCGACCCAGCGGCTCGGCTACGCTAACGAGCTCATC 1283
Qy 973 gctagattgactcacttccagtttcaagactcaagacacaaacttctactaaccacacttggactct 1032
Db 1284 GCCGCTGACCCACTGCCCTGTCCAGATGACACCACTTCCAAACCACTTTGGACTCG 1343
Qy 1033 aaccagctacttcccaattgaacgctactttgttacgctgacttctctcagcaacaact 1092
Db 1344 AGCCGGCTACTCTTCGCTCAACTCTACCTCTACCGGACCTTTCGATGACACGGC 1403
Qy 1093 atgatatctatttctctgcttgggtttggtttgtaacacggttaccagccattgttactact 1152
Db 1404 ATCACTCTCACTCTCTTGTCTTTAGGTTCTGTACCAACGGCACTAAGCGGCTATCTACCAAG 1463
Qy 1153 tctgttaactattgaagaaactgacggttactctctcttcttctgactgttccattcogct 1212
Db 1464 ACCGTGGAGATATACCCACAGACAGATGGATTCGTCTCTCTGGAGCGGTTCGCTTGTCT 1523
Qy 1213 gctagagcttacgttgaaatgatcaatgtcaagctgaaaggaacacattggttagatt 1272
Db 1524 TCGCGTTTGTACGTCGAGATGATGACAGTGTACGCGGAGCAGGCGGCTGGTCCGTGTC 1583
Qy 1273 ttggttaacgacagattgttccactgacggttggctgttgcgttgcacaaagttgggttagatt 1332
Db 1584 TTGTTTAATGATCGGTGTGCGCGCTGCATGGGTGTCCGCTTGTATGCTTTGGGAGATGT 1643
Qy 1333 aagagagacacactcgttgaaaggtttgttcttctgctagatctgtgtgtaactggctgaa 1392
Db 1644 ACCCGGATAGCTTTGTGAGGGGTTGAGCTTGTCTAGATCTCGGGGTGTGTCGGCGGAG 1703
Qy 1393 tgtttcctta 1403
Db 1704 TGTTTGTCTTA 1714

```

RESULT 8

```

US-07-923-724-7
: Sequence 7, Application US/07923724
: Patent No. 5780292
: GENERAL INFORMATION:
: APPLICANT: Nevalainen, Helena K.M.
: APPLICANT: Paloheimo, Marja T.
: APPLICANT: Miettinen-Oinonen, Arja S.K.
: APPLICANT: Torkkeli, Tuula K.
: APPLICANT: Cantrell, Michael
: APPLICANT: Piddington, Christopher S.
: APPLICANT: Rambosek, John A.
: APPLICANT: Turunen, Marja K.
: APPLICANT: Fagerström, Richard B.
: TITLE OF INVENTION: Production of Phytase Degrading Enzymes
: TITLE OF INVENTION: in Trichoderma
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESS: Sterne, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07923,724
: FILING DATE: 31-JUL-1992

```

```

: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/496,155
: FILING DATE: 19-MAR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/044,077
: FILING DATE: 29-APR-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: UK 8610600
: FILING DATE: 30-APR-1986
: ATTORNEY/AGENT INFORMATION:
: NAME: Cimbala, Michele A.
: REGISTRATION NUMBER: 33,851
: REFERENCE/DOCKET NUMBER: 1050.0240004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2363 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: both
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(404...447, 550...1906)
: US-07-923-724-7

```

```

Query Match 44.8%; Score 629.4; DB 1; Length 2363;
Best Local Similarity 65.8%; Pred. No. 4.4e-188;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;

Qy 13 gtctgtactgtccactgtccacactgttgcgttccacacacggtaccgcttgggtctct 72
Db 518 GTGGGACTACTGATCGCTGACAAATCTGTGCAGAGTCACCTCGGAGTGGCAGTCCCCGCC 577
Qy 73 cgtgtaattctcactctgtgacactgttgacggtgttaccacattgttccccagaaatt 132
Db 578 TCGAGAAATCAATCCACTTGGGATACGGTCGATCAAGGGTATCAATGCTTCTCGGAGACT 637
Qy 133 tctcacttgggtgtacacttctcacttctcttcttcttcttcttcttcttcttcttcttct 192
Db 638 TCGCATCTTTGGGGTCATACGGCGCGTCTCTCTCTGCGCAAGAAATCGGCCATCTCC 697
Qy 193 ccagacgttccagacgactgttaggttacttctgttcaagttttgtcttagacacggtgct 252
Db 698 CCGTGTGTCGCGCGGTTGCAGAGTCACTTTGCGTCAAGTCTCTCCGTCATGGAGCG 757
Qy 253 agatacccaacttctctgcctcctaaggttactctgttcttcttcttcttcttcttcttct 312
Db 758 CGGTATCCGACCGAGTCCCAAGGCAAGAAATACCTCCGCTCTCATTTGAGGAGATCCAGAG 817
Qy 313 aacgtactgtcttcaaggttaagtaagcttcttcttcttcttcttcttcttcttcttcttct 372
Db 818 AACGTGACCACTTTGATGAGAAATATGCCCTTCTCTGAGACATACAACTACAGCTTGGGT 877
Qy 373 gctgacgactgttctcattcgttgggtgaaacacaaatgggttaactctgttatttaagttctac 432
Db 878 GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGCTAGTCAACTCCGCGCATCAAGTTCTAC 937
Qy 433 agaagatacaaggttggcttagaagattgttccattcattagagcttctgttctgtgac 492
Db 938 CAGCATACGAATCGCTACCAAGGAACATCATTCCTGTTTCTCCGTTCTCTTGGCTCCAGC 997
Qy 493 agagtattgtctgtgaaagttcttcttcttcttcttcttcttcttcttcttcttcttcttct 552
Db 998 CGCTGTATCGCTTCGGGAGAGAAATTCATTTGAGGGCTTCCAGAGACCAAGCTGAAGAT 1057
Qy 553 ccaggttctcaaccacacacacacacacacacacacacacacacacacacacacacacacac 612
Db 1058 CCTGTGCCAGCGGGGCCAATCGTCGCCCAAGATCGACGTGGTCAATTTCCGAGGCCAGC 1117

```

QY 613 ggttaacacacacttggaccacggttaactgtactgttcttcaagaactctgaattaggt 672
Db 1118 TCATCCAAACAACTCTCGACCCAGGACCTGACACTGCTCTTGAAGACACGCAATTCGGCC 1177
QY 673 gaacgactgtgaagtaacttcaactgttcttgcctcagctattagatagatagaa 732
Db 1178 GATACCGTCGAAGCAATTCACCGCCACGTTTCGCCCTCCATTCGTCAACGTCGAG 1237
QY 733 gctgacttgcaggtgttaacttgaactgaacgaagacgttcttacttgatggacatgtgt 792
Db 1238 AAGACCTGTCTGCGTGACTCTCACAGACACAGAGTACCTACCTTCATGAGCATGTGC 1297
QY 793 ccattcgacactgtcgttagaacttctgaagctactgaactgaactgaactgtctgtgtttg 852
Db 1298 TCCCTTCGACACCACTTCACACGACCGTTCACCAAGCTGTCTCCCTCTCTGTGACCTG 1357
QY 853 ttcactcagcgaatgatccatacgaactacttgaagacttgggaagtactacggt 912
Db 1358 TTCACCCATGAGGAATGATCCACTAGCACTACCTCCAGTCCCTGAAAAAATACTACGGC 1417
QY 913 tacggtgtgttaacacacttgggtccagctcaaggtgtgtttcgttaacgaattgatt 972
Db 1418 CATGCGCAGGTAAACCCGCTCGGCCCGACCGGCTCGGCTACGCTACGTAAGAGCTCATC 1477
QY 973 gctagattgactcacttccagttcaagacacacacttctactaaccacacttttgactct 1032
Db 1478 GCGGCTCACCACTCCGCTGTCACGATGACACGCTCCACACACACTTGGACTCG 1537
QY 1033 aaccagacttctccacttgaacgctacttgtacgctgacttctctcaacacacact 1092
Db 1538 AACCCAGCTACTTCCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAAGGGC 1597
QY 1093 atgatatatttttctgttgggtttgtacacggttaccaggtaccagcttctactact 1152
Db 1598 ATCATCTTATCTCTGCTTTGGGTCTGTACACGCGCAATAAGCCGCTGTCTACCAAG 1657
QY 1153 tctgtgaactcttgaagaactgacggttactctgtcttctgtgactgttccatctgct 1212
Db 1658 ACCGTGGAGATATACCCACACAGATGGTCTCGTCTGCTGTGGACGTTCCGCTTGTCT 1717
QY 1213 gctagacttacttgaatgatgaatgtcaagctgtgtgtgtgacaaagtgtggttagatt 1272
Db 1718 TCGCGTCTGTAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
QY 1273 ttggttaacgacagactgttcttccatgtacacggtgtgtgtgtgtgtgtgtgtgtgt 1332
Db 1778 TTGGTTAATGATCGCGTGTCTCCGCTGATGGGTGATGGGTGATGGGTGATGGGTGAT 1837
QY 1333 aagagacgacttctgttgaaggttcttctgtcgttagatgtgtgtgtgtgtgtgtgtgt 1392
Db 1838 ACCCGGATAGCTTTGTGAGGGGGTGTGAGCTTTGTGATCTGAGGGGTGATTTGGCGGAG 1897
QY 1393 ttttctgctta 1403
Db 1898 TGTTCTGCTTA 1908

RESULT 9
US-08-609-426A-7
Sequence 7, Application US/08609426A
Patent No. 5830733
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloheimo, Marja T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Rambosek, John A.
APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
APPLICANT: Houston, Christine S.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes

TITLE OF INVENTION: in Trichoderma
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,426A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,724
FILING DATE: 31-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Grant E.
REGISTRATION NUMBER: P-41,264
REFERENCE/DOCKET NUMBER: 1050.0080001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2363 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: Join(404..447, 550..1906)
US-08-609-426A-7

Query Match 44.8%; Score 629.4; DB 2; Length 2363;
Best Local Similarity 65.8%; Pred. No. 4.4e-188;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
QY 13 gtcgtgctactgtccattgccacctgttgcgttcacacatccgggtaccgccttgggtct 72
Db 518 GTGGGACTACTGATCGTGACAACTCTGTGCAGAGTCACCTCCGAGCTGGCGCTCCCGCC 577
QY 73 cgtggttaattctcacttctgtgacactgttgacggtgtgttacaaatgtttccagaaatt 132
Db 578 TCGAGAAATCAATCCACTTTCGATACGGTTCGATCAAGGGGTATCAATGCTTCTCCGAG 637
QY 133 tctcactgtgggggtacctactctccatacttcttcttgacagacgaactgtctattct 192
Db 638 TCGCATCTTTGGGGTCAATACGCGCCCTTCTTCTCTCTGCAACCAACGATCGGCCATCTCC 697
QY 193 ccagacttccacacgactgttagatttacttctgttcttaagttttgtctagacacggtgt 252
Db 698 CCTGATGTGCGCGCGGTTGTCAGAGTCACTTTCGCTCAGGTCTCTCCGCTCATGGAGCG 757
QY 253 agataccacacttcttctgtggttaaggttactgtcttctgttgaagctattcaaaag 312
Db 758 CGGTATCCGACCGAGTCCAAAGGGCAAGAAATACTTCCGCTCTCTATTGAGGAGATCCAGCAG 817

QY 313 aacgactactgttttaagggtaagtaagctttttgaagactttacaactacaactttgggt 372
DB 818 AACGTGACCACTTTGATGGAATAATATGCTTCTGAAGACATACACTACAGCTTGGGT 877
QY 373 gctgacgactgactccattcgggtgaacacaaatggttaactctggtattagttctac 432
DB 878 GCAGATGACCTGACTCCCTCGGAGACGAGAGTAGTCAACTCCGGCATCAAGTTCTAC 937
QY 433 agaagatacaaggctttggctagaagatgttccattcattagagcttctgttctgac 492
DB 938 CAGCGATACGAATCGCTACACAGGAACATATTCGGTTCAATCCGATCCTTGGCTCCAGC 997
QY 493 agattatgttctgtgaaagattcattgaaggtttccaaatgctgaagttggtgac 552
DB 998 CGCGTATCGCTCGGCGAGAAATTCATTGAGGCTTCCAGAGCACCAGCTGAAGGAT 1057
QY 553 ccaggttctcaacacacacacaggtttccagttattcaagctgatactccagaagatcc 612
DB 1058 CTGTGTCGCCCGCGCCCAATCTCGCCCAAGATCGACTGTTCAATTCGAGGCCAGC 1117
QY 613 ggttacacacacactttggaccacggtactgttactgtcttccaaagactgtaattggt 672
DB 1118 TCATCCACACACTCTCGACCCAGGCACTGCACTCTTTGAAGACACAGCAATGGCC 1177
QY 673 gacagcttgaagtaacttcaactgtttgttgcctccagctattagagctagattggaa 732
DB 1178 GATACCGTCGAAGCAATTCACCGCCACCTTCGCCCCCTCCATTCGTCAACGCTGGAG 1237
QY 733 gctgacttgcaggtgttacttgaactgacgaagaagtggtttacttgaatgacatgtgt 792
DB 1238 AACGACCTGTGGCGTGACTCTCACAGACACAGAAAGTGACCTTACCTTGGAGATGTC 1297
QY 793 ccattgacactgctcgctagaacttctgacgactacgaatgttgcattctcattctgtg 852
DB 1298 TCCTTCGACCACTCTCCACGACGCTCGACCAAGCTGCCCCCTCTGTGACCTG 1357
QY 853 ttoactcacgaacgaatggtatccaatcagactacttgcgaagcttgggttaagtactacgt 912
DB 1358 TTCACCCATGCAATGGATCCACTAGGACTACTTCCAGTCCCTGAAATAACTACGGC 1417
QY 913 taggtgtgtaaaccaattggtcagctcagctcagctcagctcagctcagctcagctc 972
DB 1418 CATGGCGCAGGTAAACCGCTCGCCCGCCAGCCAGGCGCTGCTGCTAACGAGCTATC 1477
QY 973 gctagattgactcactctcagttcagttcagttcagttcagttcagttcagttcagtt 1032
DB 1478 GCCGCTCACCACTCGCTGCTCCAGATGACCCAGCTTCCAAACACACACCTTGGACTCG 1537
QY 1033 aacccagctacttcccaattgaacgctacttgttgcgctgacttctctcagcacaact 1092
DB 1538 AACCCAGCTACCTTCCCGCTCACTACTCTCTAGCCGACTTTCCACCATTAACGGC 1597
QY 1093 atgatactattttctgttgggtttgttgaacggttaccagccattgttctactact 1152
DB 1598 ATCATCTACTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1657
QY 1153 tctgtgaactattgaagaactgaacggttactgttcttgcgactgttccattcgtct 1212
DB 1658 ACCGTGGAGAAATACCCAGACAGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1717
QY 1213 gctagagcttgaattgaatgataatgataatgataatgataatgataatgataatg 1272
DB 1718 TCGCGCTCTAGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1777
QY 1273 ttggttaacgacagattgttccattgacggttgcgttgcgttgcgttgcgttgcgtt 1332
DB 1778 TTGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1837
QY 1333 aagaagacgacttctgtgaagattgttcttgcgttagatctgtgttaactgggtgaa 1392
DB 1838 ACCGGGATAGCTTGTGAGGGGGTTGAGCTTGTGATGATGATGATGATGATGATGATG 1897
QY 1393 tgtttcgcctta 1403

Db 1898 TGTTCTGCTTA 1908

RESULT 10

US-08-374-652C-1
Sequence 1, Application US/08374652C
Patent No. 5834286
GENERAL INFORMATION:
APPLICANT: NEVALAINEN, HELENA K.M.
APPLICANT: PALOHEIMO, MARJA T.
APPLICANT: EAGERSTROM, RICHARD B.
APPLICANT: MLETTINEN-OINONEN, ARJA S.
APPLICANT: TURUNEN, MARJA K.
APPLICANT: RAMBOSEK, JOHN A.
APPLICANT: PIDDINGTON, CHRISTOPHER S.
APPLICANT: HOUSTON, CHRISTINE S.
APPLICANT: CANTRELL, MICHAEL A.
TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
VECTORS AND METHODS FOR EXPRESSING PHYTATE DEGRADING
TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07058
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2379 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(404..447, 550..1906)
US-08-374-652C-1

Query Match 44.8%; Score 629.4; DB 2; Length 2379;
Best Local Similarity 65.8%; Pred. No. 4.5e-188;
Matches 915; Conservative 0; Mismatches 476; Indels 0; Gaps 0;
QY 13 gtcgtgactgactgactgactgactgactgactgactgactgactgactgactgactgact 72
DB 534 GTGGGACTACTGATCGCTGACAACTGTGCGAGAGTCACTTCGGAGTGGCAGTCCCGCC 593

QY	73	cggtgaattctcaactcttgacactgttgacggtgttaccaatgtttccagaaatt	132
DB	594	TCGAGAAATCAATCTCCGATACCGTGCATCAAGGATATCAATGCTTCTCCGAGACT	653
QY	133	tctcacttgggggtacactactccataactctcttttggcagacgaattctgtatttt	192
DB	654	TCGCATCTTTGGGTCATACGCGCGTCTTCTCTCTGGCAACGAATCGGCATCTCC	713
QY	193	ccagacgtttccagacgactgtagagttacttttoggttcaagtttggctctagacaggtgct	252
DB	714	CGTGTGTGCCCGCGGTTCGAGAGTCACTTGGCTCAGTCTCTCCCGTCTATGGAGCG	773
QY	253	agatacccaactctcttgctctaaagcttactctgtttgatttgaagctatttcaaaag	312
DB	774	CGGTATCCACCGAGTCAAGGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG	833
QY	313	aacgtactgttttcaagggtaagtcagcgttttcttgaagacttacaactacactttgggt	372
DB	834	AACGTGACCACCTTTGATGGAAATATGCGCTTCTCTGAAGACATACAACTACAGCTTGGGT	893
QY	373	gctgacgacttaactcaattcgggtgaaaacaaatggttaactctggtattaagttctac	432
DB	894	GCAGATGACTGACTTCCCTTCGGAGACGAGAGCTAGTCAACTTCGGCATCAAGTTCTAC	953
QY	433	agaagatacgaagcttggctagaaagattgttcatttcattagagcttctggtctctgac	492
DB	954	CAGCGATACGAATCGTCCACAAGGAACATCATTCGGTTTCATCCGATCCTCTGGCTCCAGC	1013
QY	493	agagttattgtctctgtaaaagttcaattgaaggttttccaatctgtaagtgtgctgac	552
DB	1014	CGCGTATCGCCTCCGGCGAGAAATTCATTGAGGGCTTCCAGAGCACCAAGCTGAAGGAT	1073
QY	553	ccaggttctcaaccacaccaaagcttctccagttattaacgttgatcattccagaagatcc	612
DB	1074	CTCTGTGCCACGCGGGCCCATCTGTCGCCAAGATCGACGTGGTCAATTTCGGAGGCCAGC	1133
QY	613	ggttacacacacactttggaccaggttactgtactgcttctcgaaagactctgtaattaggt	672
DB	1134	TCATCCACAACACTCTCCGACCGAGCACTGCATGTCTTTGAAGACACGCAATTTGGCC	1193
QY	673	gacgacgttgaagctaaacttcaactcttgttcctcctcaagctattagactagatggaa	732
DB	1194	GATACCTCGAAGCCAAATTTACACCGCCAGTTCGCGCCCTCCATTCGTCAACGTCGGAG	1253
QY	733	gctgacttgcaggtgttactttgactgacgaagcgtgttttacttgatggacatggt	792
DB	1254	AACGACCTGTCTGGGTGACTCTACAGACACAGAAGTGACCTACCTCATGGACATGTGC	1313
QY	793	ccattcgacactgtcgtagaacttctgacgctactgtaattgttctcaattctgtgtttg	852
DB	1314	TCCTTTGACACCATCTCCACGACCGCTCGACCAAGCTGTCCCTCTCTGTGACCTG	1373
QY	853	ttcacttcagagaaatggatccaalacactacttgcgaaagcttgggttaagctactacggt	912
DB	1374	TTCACCCATGACGAATTGGATCCACTACGACTACCTTCAGTCCCTCTGAAATACTACGGC	1433
QY	913	tacggtgtgtgtaaccatttgggtccagctcaaggtgtgtgttctgctcaagaaattgatt	972
DB	1434	CATGGCGCAGGTAACCCGTGCGGCCGACCCAGGCGTCGGCTACGCTTAACGAGTCTATC	1493
QY	973	gctagattgactcaactctccagttcaagacacacttctactaaccacacttttgaactct	1032
DB	1494	GCCCGTCTCACCACTCGCTGTCTCACGATGACACCAAGCTCCACCAACACCTTTGGATCG	1553
QY	1033	aaccagactactttccattgaacgctactttgttaagctgacttctctcacgacaaact	1092
DB	1554	AACCCAGCTACTTCCCGTCACTACTCTCTACGCGGACTTTTCCACGATTAACGGC	1613
QY	1093	atgatatctattttcttcgcttttgggtttgtacaacggttacaagccattgtctactact	1152
DB	1614	ATCATCTATCTCTTTCCTTTGGCTGTGTACAAGCGCACTTAAGCGCTCTCTACACAG	1673

Query Match 43.68; Score 612.4; DB 3; Length 1515;
Best Local Similarity 65.58; Pred. No. 7.8e-183;
Matches 895; Conservative 0; Mismatches 471; Indels 0; Caps 0;

[illegible]


```
Db 209 cggctgacaggggtatcaatgcttctcggagagacttcgcatctttggggccaatacgcgc 268
QY 158 catacttctcttggcagacgaatctgctatttctcagagcttccagagactgttagag 217
Db 269 cgtcttctctcggcaacaatacggccatctccctctgagtctcctgcooggatgccatg 328
QY 218 taacttctgtaagtttctgtagacacggtgctagatacccaacttcttctgctcta 277
Db 329 tcaacttgcgcaggttctctcccgccatggagcacggtatccgacgactccaagggca 388
QY 278 aggccttactctgctttagaagactattcaaaagaacgctactgcttccaaggttaagt 337
Db 389 agaaataactcgcctcagagagagatccagcagaacgagcaacacttgcaggggaaat 448
QY 338 acgcttcttgaagacttaacaactacacttgggtgctgacgacttgactccattgggtg 397
Db 449 atgcttcttgaacatacaactacagcctgggctggagatgactgactccctctggag 508
QY 398 aaaacaaatggttaactctggtatttaagttctacagaagaatacaaggctttggctagaa 457
Db 509 agcaggagctggtcaactcggcgctcaagttctaccagcgatagcaagaatcgctcacaagaa 568
QY 458 agattgtccattcattagagcttctggttctgacagagttattgcttctgctgaaagtt 517
Db 569 acattgtccgcttccatccgactccaggtccagcgcgtgattgctctggtgcaataaat 628
QY 518 tcattgaaggtttccaatctgctaagttggtgactgacccaggttctcaacacacacagctt 577
Db 629 tcactgaggtcttccagagcactaagtgaaagatctctcgtgcccagcccgccaatcgt 688
QY 578 ctccagttataacgtgatcattccagaagggatccggttacaacaacactttggaccacg 637
Db 689 cgcacagatcgagctggttcaattcagagggccagcagcaccatccacacactctcgatccgg 748
QY 638 gtacttactgcttccgaagactctgaattagtgtaacagcttggaagctaaacttcaactg 697
Db 749 gcacttcacaggttttcgaagatagcgaattggccggtgacatcgagccaatttccacg 808
QY 698 cttgttcgctccagctatttagagctagattggaagctgacttgcaggttacttcttga 757
Db 809 ccagcttctcctccctcattctgcaacgctctggagacgacttctgctgctgctctca 868
QY 758 ctgacgaagacgttgttacttgatgagacatgtgtccattcgacactgtcgtagaactt 817
Db 869 cggacacagaagtgactccctcattgagacatgtgtcctctcgacacactctccaccagca 928
QY 818 ctgacgctactgaattgttccattctgcttctgttctcactcagcagaatggatccaat 877
Db 929 ccgtcgacacacagctgccccctctgtgactgttccaccatgaagaatggatccaact 988
QY 878 acgactacttgaaagcttggtagtactacggttaccggtgctgtaaccattgggtc 937
Db 989 acgactactccagtcctcctgaaacataactacgcccattggcaggttaaccgctcgcc 1048
QY 938 cagctcaaggtgtgttctcgttaacgaattgattgctagattgactcaactccagttc 997
Db 1049 cgaccagggctcggtcgtacgagctatcgccgctcagctcaccactcgcgtctgctcc 1108
QY 998 agagacacactctactaaccacacttggacttaaccacagactcttccacttgaagc 1057
Db 1109 acgatgacaccgctccaaccacacattggactccaaccggtacttcccgctccaact 1168
QY 1058 ctactttagcactgacttctcagcaacacactatgatatactatttcttctgctttgg 1117
Db 1169 ccactctctatcgacttctcgtatgataacggcatcatctctactctcttcttgg 1228
QY 1118 gttgtacaacggttaccagccattgtactacttctgtgaaactatttgaagaacactg 1177
Db 1229 gctgtacaacggcacaagccgctgtcttccacgaccgcgaggagaatatacccccagaccg 1288
QY 1178 acggttactcttcttcttggactgttccattcgtgctgtagagcttaactgtgaaatgatgc 1237
Db 1289 atgggttctcactctcgtcgacggttctcttctcgtcgcgcatgtacgtcagatgatgc 1348
```

```
QY 1238 aatgtcaagctgaaaaggaaaccattggttagagtttggtaacgacagagttgttccat 1297
Db 1349 aatgcccagtcgcagcagcagcagccttggctcgcgtcttggtaatgctgtgttgcgc 1408
QY 1298 tgcacaggttctgctgttgacaagcttgggttagatgtgaagagagacgactcgttgaaggtt 1357
Db 1409 tgcagctgctcgcgttggatgtcttgggaagatgtacgcgggatgcttctggaaggggt 1468
QY 1358 tgccttctgctagatcgttggtaactggctggaatgttctgcctta 1403
Db 1469 tgaagcttccagatcgtcgcgttgattggggggagtgcttccgctta 1514

RESULT 12
US-09-155-855-4
; Sequence 4, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155.855
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ. ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1332)
US-09-155-855-4
```

Query Match 43.5%; Score 610.2; DB 3; Length 1332;
Best Local Similarity 66.2%; Pred. No. 3.6e-182;
Matches 879; Conservative 0; Mismatches 448; Indels 0; Gaps 0;

```
QY 77 gtaattctcactctgtgacactgttgacggtggttaccatgtttcccaagaattctc 136
Db 5 gaatacaatccacttgcgtacgcgtcgcaggggtatcaatgcttctcgagacttcgc 64
QY 137 acttgtgggttacctctccactctctcttcttggcagacgaatctgctatttctccag 196
Db 65 atcttggggcacaatacgcgcgttcttctctggaacaaatcgccactctccctg 124
QY 197 acgttccagacgactgttagagttacttctgcttcaagtttcttagacaggtgtagat 256
Db 125 atgtcttcgcggagtcgcatgctcacttctcccgaggttctctccgcatgggacggt 184
QY 257 acccaacttctctgcgtctcaagcttactctgcttcttggattgaagctattcaaaaagc 316
Db 185 atccgaccgactccaaggcgaagaaatactcctctctcgcggaggatccagagaagc 244
QY 317 ctactgcttcaagggttagtacgcttcttgaagacttacaactacacttttgggtgctg 376
Db 245 cgacaaccttcgagggtgaataatgcttctctgaacatacaactacagcctggcgcg 304
QY 377 acgacttgaactcattcgttgaaacaaaatggttaactctgttataagttctacagaa 436
Db 305 atgacttgaactcctctcgtgagagcagagctgtgtaactccgctcaggttctaccagc 364
QY 437 gatacaaggtcttggctagaaagattgttccattcattagagcttctggttctgacagag 496
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:25 ; Search time 75.85 Seconds
(without alignments)
814.587 Million cell updates/sec

Title: US-09-488-265-29

Perfect score: 2462

Sequence: 1 MGFFVLLSTATLFGSTSGT.....DFVEGLSTARSGNNAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: .425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1868	75.9	467	3 Q9U27	Q9U27 aspergillus
2	1867	75.8	465	3 O00092	O00092 aspergillus
3	1860	75.5	467	3 Q93838	Q93838 aspergillus
4	1850	75.1	467	3 Q9HEQ0	Q9HEQ0 aspergillus
5	1839	74.7	466	3 O00100	O00100 aspergillus
6	1761	71.5	466	3 O00085	O00085 aspergillus
7	1719	69.8	466	3 O00096	O00096 aspergillus
8	1313.5	53.4	487	3 O00107	O00107 thielavia h
9	344.5	14.0	469	3 Q9Y846	Q9Y846 kluyveromyc
10	328.5	13.3	442	3 Q74677	Q74677 pichia angu
11	297.5	12.1	463	3 O60172	O60172 schizosacch
12	203	8.2	460	5 Q9VW72	Q9VW72 drosophila
13	192.5	7.8	467	5 O96421	O96421 drosophila
14	189	7.7	451	11 Q35217	Q35217 rattus norv
15	186	7.6	481	11 Q9Z2L6	Q9Z2L6 mus musculus
16	175.5	7.1	453	5 O96420	O96420 drosophila
17	175.5	7.1	487	4 Q9UNW1	Q9UNW1 homo sapien
18	174.5	7.1	453	5 O96438	O96438 drosophila
19	174.5	7.1	487	4 O95172	O95172 homo sapien

20	173.5	7.0	487	4 Q9UCA3	Q9UGA3 homo sapien
21	170.5	6.9	449	13 Q92170	Q92170 gallus gall
22	145.5	5.9	468	10 O04509	O04509 arabidopsis
23	136.5	5.5	198	3 Q9UTX1	Q9UTX1 schizosacch
24	136	5.5	274	11 Q9LUD5	Q9JJD5 mus musculu
25	131.5	5.3	374	11 Q9JMG5	Q9JMG5 mus musculu
26	129.5	5.3	381	11 Q9CXG5	Q9CXG5 mus musculu
27	119	4.8	449	5 Q19076	Q19076 caenorhabdi
28	117	4.8	683	5 O00838	O00838 leishmania
29	117	4.8	707	5 O00839	O00839 leishmania
30	114.5	4.7	513	5 O06334	O06334 comamonas t
31	113	4.6	827	2 Q47871	Q47871 eubacterium
32	110.5	4.5	421	4 Q9UIG6	Q9UIG6 homo sapien
33	110.5	4.5	428	4 Q9NPH0	Q9NPH0 homo sapien
34	107.5	4.4	1013	14 Q9ILX9	Q9ILX9 retroperito
35	106.5	4.3	380	5 Q22525	Q22525 caenorhabdi
36	106	4.3	853	14 Q93123	Q93123 human calic
37	104	4.2	354	5 Q19709	Q19709 caenorhabdi
38	103	4.2	1225	5 Q20330	Q20330 caenorhabdi
39	102	4.1	452	5 Q19175	Q19175 caenorhabdi
40	101.5	4.1	381	11 Q9QXH7	Q9QXH7 mus musculu
41	101.5	4.1	516	5 Q25327	Q25327 leishmania
42	101.5	4.1	1723	2 Q72194	Q72194 porphyromon
43	100.5	4.1	513	2 Q24719	Q24719 comamonas t
44	100.5	4.1	730	5 Q20826	Q20826 caenorhabdi
45	99.5	4.0	758	4 Q9F2C1	Q9F2C1 homo sapien

ALIGNMENTS

RESULT 1
Q9U27 PRELIMINARY; PRT; 467 AA.
ID Q9U27
AC Q9U27;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (BC 3.1.3.8).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
RX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RA Hongning W., Qi W., Jing X.;
RT "PCR, cloning and characterization of the phytase (phyA) gene of
RT Aspergillus niger (China Strain).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218813; AAF25481.1;
DR HSP; P34752; IHP.
DR InterPro; IPR000560;
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Signal; Lyase; Hydrolase.
FT SIGNAL; 19 POTENTIAL.
SQ SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 75.9%; Score 1868; DB 3; Length 467;
Best Local Similarity 75.2%; Pred. No. 1.3e-141;
Matches 351; Conservative 43; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGFFVLLSTATLFGSTSGTALGFRGNHSCDITVDGGYQCFPEISHLWGTSPYFSLADE 60
Db 1 MGVSALLPLVLLSGVTSGLAVPASRNQSTCDTVDGGYQCFSETSHLWGWAPFESLANE 60

QY 61 SAISDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALEIAIQNATFAKGYAFLKTYN 120
Db 61 SAISDPVAGCRVTFVQVLSRHGARYPTSSASKAYSALEIAIQNATFAKGYAFLKTYN 120

QY 121 YTLGADDLTPGEMQVNSGKIFVRRYKALARKIVPPIFRASGSDRVIASAEKFTGQSA 180

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543EDC265C CRC64;

Query Match 75.5%; Score 1860; DB 3; Length 467;
 Best Local Similarity 74.1%; Pred. No. 5.8e-141;
 Matches 346; Conservative 48; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSDTVDDGGYQCFPEISHLWGTYSPYFSLADE 60
 Db 1 MGVSAVLLPLLYLLSGVTSGLAVPASRNQSTCDTVDDGGYQCFSETSHLWGYPFFSLANK 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFLLKTYN 120
 Db 61 SAISPDVAGCHVTFAQVLSRHGARYPTDSKGYKYSALIEEIQONATTFEGKYAFLLKTYN 120
 QY 121 YTLGADDLTPFGENQMVNSGKIFRYRKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 Db 121 YSLGADDLTPFGEQELVNSGVKQYRYSLETRNIVPFIRSSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGSOPHOASPVINVIIEGSGYNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 KLKDPRAQPGOSSPKIDVISEASTSNNTLDPGTCTVFEDSELADDEANFTATFVPSIR 240
 QY 241 ARLEADLPGLVTLTDVYVLMDCPFDVTARTSDATLSPECALFTDHEWIDYQLQSLG 300
 Db 241 QRLNDLSGVSLTDETVYVLMDCSFDTISTSTVDTKLSPFCDLFTHEWINYDIQLSLN 300
 QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYYGAGNPLGTPQGVYANELIARLTHSPVHDDTSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNMTISIFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTAENITQTDGFSARTVPFASRMVYEMMQCSEQEP 420
 QY 421 LVRVLNDRVPLHGCADVKLGRCRDDFVGLSFARSGGNWAECEFA 467
 Db 421 LVRVLNDRVPLHGPCVDALGRCTRDSFVKGLSFARSGGDWAECEFA 467

RESULT 4

Q9HE00 PRELIMINARY; PRT; 467 AA.
 AC Q9HE00;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PHYTASE.
 OS Aspergillus ficum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 RN NCBI_TaxID=5058;
 RP SEQUENCE FROM N.A.
 RA Zhang L., An L., Wang Y., Yuan X.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY013315; AAG40885.1;
 SQ SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match 75.1%; Score 1850; DB 3; Length 467;
 Best Local Similarity 74.1%; Pred. No. 3.7e-140;
 Matches 346; Conservative 48; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSDTVDDGGYQCFPEISHLWGTYSPYFSLADE 60
 Db 1 MGVSAVLLPLLYLLSGVTSGLAVPASRNQSTCDTVDDGGYQCFSETSHLWGYPFFSLANK 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFLLKTYN 120
 Db 61 SAISPDVAGCHVTFAQVLSRHGARYPTDSKGYKYSALIEEIQONATTFEGKYAFLLKTYN 120

QY 121 YTLGADDLTPFGENQMVNSGKIFRYRKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 Db 121 YSLGADDLTPFGEQELVNSGVKQYRYSLETRNIVPFIRSSGSRVIAAGNKFIEGFQST 180
 QY 181 KLADPGSOPHOASPVINVIIEGSGYNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 KLKDPRAQPGOSSPKIDVISEASTSNNTLDPGTCTVFEDSELADDEANFTATFVPSIR 240
 QY 241 ARLEADLPGLVTLTDVYVLMDCPFDVTARTSDATLSPECALFTDHEWIDYQLQSLG 300
 Db 241 QRLNDLSGVSLTDETVYVLMDCSFDTISTSTVDTKLSPFCDLFTHEWINYDIQLSLN 300
 QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYYGAGNPLGTPQGVYANELIARLTHSPVHDDTSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNMTISIFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTAENITQTDGFSARTVPFASRMVYEMMQCSEQEP 420
 QY 421 LVRVLNDRVPLHGCADVKLGRCRDDFVGLSFARSGGNWAECEFA 467
 Db 421 LVRVLNDRVPLHGPCVDALGRCTRDSFVKGLSFARSGGDWAECEFA 467

RESULT 5
 Q00100 PRELIMINARY; PRT; 466 AA.
 AC Q00100;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DE 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE PHOSPHOHYDROLASE).
 DE Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 RN NCBI_TaxID=33178;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CBS-116.46;
 RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B., Broger C., van Loon A.P.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U60412; AAB58465.1;
 DR HSP; P34752; 1IHP.
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 466
 FT ACT_SITE 82 82
 FT ACT_SITE 83 83
 FT ACT_SITE 361 361
 FT DISULFID 31 40
 FT DISULFID 71 414
 FT DISULFID 215 465
 FT DISULFID 264 282
 FT DISULFID 436 444
 FT CARBOHYD 27 27
 FT CARBOHYD 105 105
 FT CARBOHYD 120 120

FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51055 MW; F2AECECIAF7C22C4 CRC64;

Query Match 74.7%; Score 1839; DB 3; Length 466;
 Best Local Similarity 73.0%; Pred. No. 2.8e-139;
 Matches 340; Conservative 49; Mismatches 77; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTSGTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60
 DB 1 MGFFVLLSIATLFGSTSGTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60

QY 61 SAISPDVDDCVTFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAPKGYAFLKTYN 120
 DB 61 SPFPLDVPDCHITFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAPKGYAFLKTYN 120

QY 121 YTLGADDLTPFGENOMVNSGKIFRYKALARKIVPTIRASGSDRVIASAEKTEGFQSA 180
 DB 121 YSMGSENLNPFGRNQLDGLGAQFYRYDITLRIINPFVRAADSSRVHESAEKTEGFQSA 180

QY 181 KLADPGSQHQAQSPVNIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 ROGDDHANPHQSPRVVDVAIPEGSAIYNTLEHSLCTAFESSTVGDDAVANFTALFAPAIR 240

QY 241 ARLEADLPVGLTDEDVYVIMDMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 DB 241 KRLEADLPVGLTDEDVYVIMDMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300

QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATPLNATLYADFS 360

QY 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420
 DB 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECF 466
 DB 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECF 466

RESULT 6
 ID 000085 PRELIMINARY; PRT; 466 AA.
 AC 000085;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE, 3-
 DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 GN PHA.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 ON NCBI_TaxID=33178;
 RX STRAIN=9A1;
 RX MEDLINE=97177792; PubMed=9025238;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
 RT "The phytase subfamily of histidine acid phosphatases: isolation of
 RT genes for two novel phytases from the fungi *Aspergillus terreus* and
 RT *Myceliophthora thermophila*.";
 RL Microbiology 143:245-252(1997).
 CC -[- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -[- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O -> D-MYO-
 CC -INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

CC -[- SUBCELLULAR LOCATION: SECRETED.
 CC -[- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 CC 2.5 TO 7.5, WITH MAXIMAL ACTIVITY AT PH 5.5. ALSO ACCEPT 4-
 CC NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO
 CC MORE ACIDIC PH VALUES.
 CC -[- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U59805; AAB52507.1; -;
 DR HSSP: P34752; LIHP.
 DR InterPro: IPR000560; -;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 466 REQUIRED FOR BINDING SUBSTRATE (BY
 FT ACT_SITE 82 82 SIMILARITY).
 FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 31 40 BY SIMILARITY.
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51093 MW; 21DCB559C96AB66 CRC64;

Query Match 71.5%; Score 1761; DB 3; Length 466;

Best Local Similarity 69.3%; Pred. No. 5.1e-133;

Matches 324; Conservative 58; Mismatches 84; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTSGTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60
 DB 1 MGFFVLLSIATLFGSTSGTALPGRNHSCDVTGGYQCPEISHLWGTSPYFSLADE 60

QY 61 SAISPDVDDCVTFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAPKGYAFLKTYN 120
 DB 61 SPFPLDVPDCHITFVQVLSRHGARYPTSSAKAYSALIEAIQKNATAPKGYAFLKTYN 120

QY 121 YTLGADDLTPGENOMVNSGKIFRYKALARKIVPTIRASGSDRVIASAEKTEGFQSA 180
 DB 121 YLSDSEELTPGRNQLDGLGAQFYRYNALTIRHINPFVRAATDSRVHESAEKTEGFQSA 180

QY 181 KLADPGSQHQAQSPVNIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 RQDDHANPHQSPRVVDVAIPEGSAIYNTLEHSLCTAFESSTVGDDAVANFTALFAPAIR 240

QY 241 ARLEADLPVGLTDEDVYVIMDMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 DB 241 QRLEADLPVGLTDEDVYVIMDMCPFTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300

QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGGNPLGPAQGVGFANELLARLTHSPVODHTSTNNHTLDSNPATPLNATLYADFS 360

QY 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420
 DB 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCAKEP 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECF 466
 DB 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECF 466

RESULT 7

000096

ID 000096

PRELIMINARY;

PRT; 466 AA.

RC RX MEDLINE=98028656; PubMed=9359836;
 RA Craxton A., Caffrey J.J., Burkhardt W., Safrany S.T., Shears S.B.;
 RT "Molecular cloning and expression of a rat hepatic multiple inositol
 RL polyphosphate phosphatase";
 RN Biochem. J. 328:75-81(1997).
 RN [2]
 RP CATALYTIC ACTIVITY.
 RX MEDLINE=91358435; PubMed=1653239;
 RA Nogimori K., Hughes P.J., Glennon M.C., Hodgson M.E., Putney J.W. Jr.,
 RA Shears S.B.;
 RT "Purification of an inositol (1,3,4,5)-tetrakisphosphate 3-phosphatase
 RT activity from rat liver and the evaluation of its substrate
 RT specificity";
 RL J. Biol. Chem. 266:16499-16506(1991).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INOSITOL 1,3,4,5-
 CC TETRAPHOSPHATE TO INOSITOL 1,4,5-TRIPHOSPHATE AND INOSITOL
 CC 1,3,4,5,6-PENTAKISPHOSPHATE TO INOSITOL 1,4,5,6-TETRAKISPHOSPHATE
 CC AND THEN TO INOSITOL 1,4,5-TRISPHOSPHATE. DEPHOSPHORYLATES
 CC INOSITOL HEXAKISPHOSPHATE WITHOUT SPECIFICITY TOWARDS A PARTICULAR
 CC PHOSPHATE GROUP.
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY AND LIVER.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL, AF012714; AAC53453.1; -
 DR InterPro; IPR000560; -
 DR InterPro; IPR000886; -
 DR Pfam; PF00328; acid.phosphat; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; FALSE_NEG.
 KW Hydrolase; Glycoprotein; Phosphorylation; Multifunctional enzyme.
 FT ACT_SITE 59
 FT MOD_RES 65
 FT MOD_RES 146
 FT MOD_RES 176
 FT MOD_RES 201
 FT MOD_RES 201
 FT MOD_RES 218
 FT MOD_RES 391
 FT MOD_RES 412
 FT MOD_RES 435
 FT MOD_RES 447
 FT MOD_RES 447
 FT CARBOHYD 206
 FT CARBOHYD 445
 FT VARIANT 3
 FT CONFLICT 327
 FT SITE 448
 FT SEQUENCE 451 AA; 51592 MW; EB1C05512A03020B CRC64;

Query Match 7.7%; Score 189; DB 11; Length 451;
 Best Local Similarity 20.7%; Pred. No. 5.5e-07;
 Matches 89; Conservative 66; Mismatches 193; Indels 82; Gaps 16;
 QY 52 SPYFS-----LAESAISPD---VPDCCRVTFVQVLSRHGARYPTSSASKAYS 96
 DB 14 SPYFGTKTRYEDVNPWLLGDPVAPRRDPPELLAGTCTPVQLVALIRHGTTRYPTTKQIRKLR 73
 QY 97 ALIEAIOKNATAFKGYAFKLTNTYTLGADLLTPFGENQMVNSGKIFRYRYKALARKIVP 156
 DB 74 QLOGLLQTRSVSDGGRVAALDQWPLWYDD---WMDGOLVERGRQDMQLALRLAALFP 130
 QY 157 --FIR-----ASGSDRVIASAEKFIKGFQSAKLADPGSQPHQAS-----APAIRARLE-----ADL 199
 DB 131 DLFCRENGRLRLITSSKRCVSSAAFLQGLW---QHYHGLPPDPVSDMECDPPRVN-- 186
 QY 200 IPEGSGYNTLTHGCTAFEDSEIGDDVEANFTALF-----APAIRARLE-----ADL 247
 DB 187 ----DKLMRFDDH--CEKEL-----TEVERNATLYHVEAFKTPGEMQTVLKKVAATLQV 235
 QY 248 PGVTLTDEDVYVLMDCMPPFTVARTSDATLSPPCALFTHDEWIQDYDYLQSLGKYGYGA 307
 DB 236 PVNNLNADLIQVAFFTCFSFDIAIQGVH-----SPWCDVFDVDDAKVLEYLNDLKQYWKRSY 291

QY 308 GNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMIS 367
 DB 292 GYAINSRSSCNLFQDIFLHLDKAVEQKORSQ-----PVSSVILQFGHAETLLP 340
 QY 368 IFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCQAEKEP---LVR 423
 DB 341 LLSLMGVFKDKPLTAYNEEQVHRE-FRSGHIVPVASNLIFVLYHCEDAQTPOEKFQIQ 399
 QY 424 VLVNDRVVPL 433
 DB 400 MLLNEKVLP 409
 RESULT 15
 Q922L6 PRELIMINARY; PRT; 481 AA.
 ID Q922L6
 AC Q922L6;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
 GN MINPPI OR MIPP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AF046908; AAC02434.1; -
 DR MGD; MGI:1336159; Minppl.
 DR InterPro; IPR000560; -
 DR InterPro; IPR000886; -
 DR Pfam; PF00328; acid.phosphat; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
 SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;

Query Match 7.6%; Score 186; DB 11; Length 481;
 Best Local Similarity 21.0%; Pred. No. 1e-06;
 Matches 99; Conservative 66; Mismatches 206; Indels 100; Gaps 18;
 QY 11 ATLFGSTGALGPRGNHSCDVTVDGYQCPEISHLWGTSPYFS-----LA 58
 DB 21 AALLSSFARCSLPGRGD-----PVASVL-----SPYFGTKTRYEDANPWLIV 62
 QY 59 DESAISPD---VPDCCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAF 115
 DB 63 DPVAPRRDPPELLAGTCTPVQLVALIRHGTTRYPTTKQIRKQLQGLLQTRSDRGSSQA 122
 QY 116 LKTYNTLGCADLLTPFGENQMVNSGKIFRYRYKALARKIVP--FIR-----ASGSD 164
 DB 123 AALAEWPLWYGD---WMDGOLVERGRQDMQLALRLAALFPDLFSRENYDLRLITSSKH 179
 QY 165 RVIASAEKFIKGFQSAKLADPGSQPHQAS-----PVINVIPEGSGYNTLTHGCTAF 218
 DB 180 RCVDSAAFLQGLW---QHYHGLPPDPVSDMECDPPRVN-----DKLMRFDDH--CEKF 229
 QY 219 EDESELGDDVEANFTALF-----APAIRARLE-----ADLPGVTLTDEDVYVLMDCMPF 266
 DB 230 L-----TDVERNATLYHVEAFKTPGEMQVKLVKVAATLQVPMNSLNADLIQVAFFTCF 284
 QY 267 DIVARTSDATLSPPCALFTHDEWIQDYDYLQSLGKYGYGAGNPLGPAQGVGFANELIAR 326
 DB 285 DLAIKGVH-----SPWCDVFDVDDAKVLEYLNDLKQYWKRSYGYTINSSSCNLFIDFLH 340
 QY 327 LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMISIFFALGLYNGTKPLSTTSV 386
 DB 341 LDKAVEQKORSQ-----PVSSVILQFGHAETLLPILLSLMGVFKDKEPTAYNF 389

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:28 ; Search time 48.78 Seconds
(without alignments)
729.264 Million cell updates/sec

Title: US-09-488-265-29

Perfect score: 2462

Sequence: 1 MGVFVLLSIATLFGSTSGT.....DFVEGLSFARSGNNAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	76.3	467	1 JN0889	3-phytase (EC 3.1.1)
2	1875	76.2	467	1 JN0656	3-phytase (EC 3.1.1)
3	1814	73.7	441	1 JN0482	3-phytase (EC 3.1.1)
4	395	16.0	467	2 S53476	acid phosphatase (
5	395	16.0	467	2 S48996	acid phosphatase (
6	392	15.9	467	1 PABYC	acid phosphatase (
7	380	15.4	467	1 PABYC	acid phosphatase (
8	355	14.4	468	2 JC4285	acid phosphatase (
9	354	14.4	468	2 S52495	acid phosphatase (
10	347.5	14.1	479	1 JN0890	acid phosphatase (
11	346.5	14.1	479	1 JN0890	acid phosphatase (
12	297.5	12.1	463	2 T39929	3-phytase (EC 3.1.1)
13	297	12.1	453	1 A25326	thiamin-repressibl
14	278.5	11.3	463	2 S14119	acid phosphatase (
15	149	6.1	465	2 JE0369	acid phosphatase (
16	145.5	5.9	468	2 A86233	histidine acid pho
17	130	5.3	755	2 T19118	hypothetical prote
18	124.5	5.1	421	2 S14742	acid phosphatase h
19	119	4.8	449	2 T15933	acid phosphatase h
20	114.5	4.7	423	2 A33395	hypothetical prote
21	112	4.5	721	2 T27570	hypothetical prote
22	106.5	4.3	380	2 T16883	hypothetical prote
23	104	4.2	354	2 T21241	hypothetical prote
24	103	4.2	1225	2 T16346	hypothetical prote
25	102.5	4.2	423	1 S06167	hypothetical prote
26	102	4.1	452	2 T20556	acid phosphatase (
27	101.5	4.1	680	2 S63698	hypothetical prote
28	101	4.1	563	2 S33957	NADPH--ferrihemop
29	100.5	4.1	730	2 T16455	coat protein gamma
					hypothetical prote

30 98.5 4.0 381 2 JH0152 acid phosphatase (

31 98.5 4.0 642 2 E69144 probable formate C

32 98.5 4.0 680 2 S63895 NADPH--ferrihemop

33 97.5 4.0 715 4 TVMSMY transforming prote

34 96.5 3.9 275 2 S09774 hypothetical prote

35 96.5 3.9 386 1 JH0610 acid phosphatase (

36 96.5 3.9 1433 2 A71444 acid phosphatase (

37 96 3.9 413 2 T18945 hypothetical prote

38 95.5 3.9 473 2 G84312 glycine dehydrogen

39 95.5 3.9 680 2 A37890 NADPH--ferrihemop

40 95.5 3.9 845 2 T17291 hypothetical prote

41 95.5 3.9 1732 2 T30836 lysine-specific cy

42 95.5 3.9 4924 2 T50176 probable peptide s

43 95 3.9 605 2 JH0638 alpha-amylase (EC

44 94.5 3.8 580 1 WMBEWS UL25 protein - hum

45 94.5 3.8 602 2 A36715 exo-poly-alpha-gal

ALIGNMENTS

RESULT 1

JN0889

3-phytase (EC 3.1.1.3.8) A precursor - Aspergillus awamori

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phyA protein

C:Species: Aspergillus awamori

C>Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0889

R:PIDdington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0889

A:Molecule type: DNA

A:Residues: 1-467 <PID>

A:Cross-references: GB:I02421; NID:q166518; PIDN:AAAL6898.1; PID:g166519

A:Experimental source: strain ALK0243

A>Note: part of the sequence, including the amino end of the mature protein, was conf

C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic

C:Genetics:

A:Gene: phyA

A:Introns: 15/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-467/Product: 3-phytase A #status experimental <MAT>

F:27.59,105,120,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #

F:81.361/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 76.3%; Score 1879; DB 1; Length 467;

Best Local Similarity 75.6%; Pred. No. 1.4e-143;

Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTSGTALPGRNHSCDVTDDGGVQCQFPEISHLWGTVSFVSLADE 60

Db 1 MGVSALLPLLYLLAGVTSGTGLAVPASRNQSTCDTVDQGYQCFSETSHLWGQVAFPSFLANE 60

QY 61 SAISPDPVDDCRVTFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120

Db 61 SAISPDPVAGCRVTFQVLSRHGARYPTESKGYKYSALIEIQNNVTTFDKGYAFLKTYN 120

QY 121 YTLGADDTTPGQNMVNGIKFYRRYKALARKIVPFIRASGSDRVIASAEKFTIEGQSA 180

Db 121 YSLGADDTTPFGEOLVNSGKIFQYRSLTRNIIPIFIRSSGSRVITASGEKFTIEGQST 180

QY 181 KLADPGSOPHOASPVINIIPEGSGYNNLTDLHCTAFEDSELGDDVEANFTALFAPAIR 240

Db 181 KLKDPAPQGGSSPKIDVIVSEASSNNLTDPGCTCTFEDELADTVEANFTATFAPSIR 240

QY 241 ARLEADLPGVTLTDEDDVYVLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDLOSGLG 300

Db 241 ORLENDLSGVTLTDEVTYLLMDMCSFDITSTVDTKLSPFCDLTHDEWHYDLSQK 300
QY 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
QY 361 HNTMISIFALGLYNGTKPLSTSVESIEETDGYASAWTVPFAARAYVEMMQCAQEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTSVESIEETDGYASAWTVPFAARAYVEMMQCAQEP 420
QY 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467
Db 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467
RESULT 2
JN0656
3-phytase (EC 3.1.3.8) A precursor - Aspergillus niger
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phvA protein
C:Species: Aspergillus niger
C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.
A.M.J.J.
Gene 127, 87-94, 1993
A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phvA)
A:Reference number: JN0656; MUID:93252284
A:Accession: JN0656
A:Molecule type: DNA
A:Residues: 1-467 <VAN>
A:Cross-references: GB:216414; NID:q2392; PIDN:CAA78904.1; PID:g2393
A:Experimental source: strain NRRL3135
A:Note: Parts of the sequence, including the amino end of the mature protein, were confi
C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and ino
C:Genetics:
A:Gene: phvA
A:Introns: 15/2
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-467/Product: 3-phytase A #status experimental <MAT>
F:27,59,105,120,207,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #sta
F:81,361/Active site: Arg, His #status predicted
F:82/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 76.2%; Score 1875; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 3e-143;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;
QY 1 MGVEVVLSTATLFGSTGALGPRGNHSCDFVDGQYOCFFPEISHLWGTSPYFSLADE 60
Db 1 MGVSALLPLLYLSGVTSLGAVPASRNQSCDVTGQYOCFFSETSHLWQYAPFFSLANE 60
QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKYASALIEATOKNATAPKGYAFELKTYN 120
Db 61 SVISPEVPACRVTFVQVLSRHGARYPTSSAKYASALIEATOKNATAPKGYAFELKTYN 120
QY 121 YTLGADLTFFGENOMVNSGKIFRYKALARKIVPFIASGSDRVIASAEKFIQEGFQSA 180
Db 121 YSLGADLTFFGEQVNSGKIFRYKALARKIVPFIASGSDRVIASAEKFIQEGFQST 180
QY 181 KLADPGSOPHOASPVNIVIPESGYNNTLDHGCTCTAFEDSELGDDVEANFTALFAPATR 240
Db 181 KLKDPRAQPGQSPKIDVLISEASSNNTLDPCCTVFEDSELADTVEANFTALFAPATR 240
QY 241 ARLEADLPVGLTDEDDVYVLMDCPFTVARTSDATSELPFCALFTHDEWIDYDLSQK 300
Db 241 ORLENDLSGVTLTDEVTYLLMDMCSFDITSTVDTKLSPFCDLTHDEWHYDLSQK 300
QY 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYYGAGNPLGPAQGVAFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
QY 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467
Db 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467

QY 361 HNTMISIFALGLYNGTKPLSTSVESIEETDGYASAWTVPFAARAYVEMMQCAQEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTSVESIEETDGYASAWTVPFAARAYVEMMQCAQEP 420
QY 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467
Db 421 LVRVLNDRVPLHGCAGVADKLGRCRDRDFEGLSFARSGGNWAECEFA 467
RESULT 3
JN0482
3-phytase (EC 3.1.3.8) A - Aspergillus ficum
N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phvA protein
C:Species: Aspergillus ficum
C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
R:Ullah, A.H.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 192, 747-753, 1993
A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemi
A:Reference number: JN0482; MUID:93249451
A:Accession: JN0482
A:Molecule type: protein
A:Residues: 1-441 <ULL>
A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
Biochem. Biophys. Res. Commun. 178, 45-53, 1991
A:Title: Cyclohexanedione modification of arginine at the active site of Aspergillus
A:Reference number: PN0023; MUID:91298982
A:Accession: PN0023
A:Molecule type: protein
A:Residues: 48-70 <UL2>
C:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
C:Superfamily: yeast acid phosphatase
C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydrate (Asn) (covalent) #sta
F:58,338/Active site: Arg, His #status predicted
F:59/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 73.7%; Score 1814; DB 1; Length 441;
Best Local Similarity 77.2%; Pred. No. 2.3e-138;
Matches 338; Conservative 39; Mismatches 61; Indels 0; Gaps 0;
QY 27 NSHSCDVTGQYOCFFPEISHLWGTSPYFSLADESALSPVDDCRVTFVQVLSRHGARY 86
Db 4 NOSCDVTGQYOCFFSETSHLWQYAPFFSLANESVISPEVPACRVTFVQVLSRHGARY 63
QY 87 PTSASAKYASALIEATOKNATAPKGYAFELKTYNLTGADLTFFGENOMVNSGKIFRYR 146
Db 64 PTDGKGYKYSALIEATOKNATAPKGYAFELKTYNLTGADLTFFGEQVNSGKIFRYR 123
QY 147 YKALARKIVPFIASGSDRVIASAEKFIQEGFQSAKLADPGSOPHOASPVNIVIPESGY 206
Db 124 YESLTRLNVPFISSGSDRVIASCKKFIQEGFQSTKLADPGSOPHOASPVNIVIPESGY 183
QY 207 NNTLDHGCTCTAFEDSELGDDVEANFTALFAPATRLEADLPVGLTDEDDVYVLMDCP 266
Db 184 NNTLDHGCTCTAFEDSELADTVEANFTALFAPATRLEADLPVGLTDEDDVYVLMDCSF 243
QY 267 DTVARTSDATSELPFCALFTHDEWIDYDLSQKYYGYGAGNPLGPAQGVAFANELIAR 326
Db 244 DTISTSTVDTKLSPFCDLTHDEWIDYDLSQKYYGYGAGNPLGPAQGVAFANELIAR 303
QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMISIFALGLYNGTKPLSTTV 386
Db 304 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMISIFALGLYNGTKPLSTTV 363
QY 387 ESIEETDGYASAWTVPFAARAYVEMMQCAQEPFLVRVLNDRVPLHGCAGVADKLGRCR 446
Db 364 ENITDGTGSSAWTVPFAARAYVEMMQCAQEPFLVRVLNDRVPLHGCAGVADKLGRCR 423
QY 447 DDFVEGLSFARSGGNWAE 464

```
Db      ||||| : | : : | | | | | :  
399 NDVVPITCSTGPGFSCSEINFDYDAEKR VAGTDFLK 436  
Dl
```

RESULT

S48996

acid phosphatase (EC 3.1.3.2). PHO12 - yeast (*Saccharomyces cerevisiae*).
N; Alternate names: protein YHR215w
C; Species: *Saccharomyces cerevisiae*
C; Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999
C; Accession: S48996; S59659
R; Macri, C.
submitted to the EMBL Data Library, February 1994
A; Description: The sequence of *S. cerevisiae* cosmid 9177.

```
db      | 21 | PLGLSDIDIGTQTETFPFL---GGSGPYYSFPGDYGISRLDPESCEMQVMVGRHG 76
ov      | 84 | ARYPTSSASKAYSALTEAIOKNATAPKGKYAFIK-----TVNVTI-----CADDY 128
```

OV 129 TPF-GENOMVNSGTFYRRYKALARKTVPF-TRASGSDRVTASAEKFTEGEOSAKIADPG 186

Db 137 NPYTGEMNAKRHARDEFLAQGYMVENQTSFAVFTSNSNRCHDTAQYFIDG----LGDKF 191

QY 107 3QFIHQ3FVINVITFE93GINNIEDHGICIAFEDSEGGDDVEANTALTAPAIRARLEAD 240

SECRET

[illegible]

QY 307 AGNPLGPAOGVGFANELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFSDNTMI 366

DB 301 PGYDVVRVSGANLFNASVKLLKESEVQDQ-----KVWLSFTHDIL 342

[illegible]

OV 427 NDRVPLHGCAVDKLGRCCKRDDF-----VEGLSFAR 457

Db 399 NDAVVPIETCSTGPGFSCEINDFYGAEKRVAGTDFLK 436

RESULT 6

acid phosphatase (EC 3.1.3.2) repressible, precursor - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: acid phosphatase PHO5; protein YBR0814; protein YBR093c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 19-Feb-1984 #sequence revision 30-Sep-1991 #text change 12-Nov-1999
 C:Accession: S05795; S48260; S45961; A00777; A38793; S41855; B25241; A25367; A27
 R:Baizha, W.; Meyhack, B.; Rudolph, H.; Schweingruber, A.M.; Hinnen, A.
 Nucleic Acids Res. 12, 7721-7739, 1984
 A:Title: Structural analysis of the two tandemly repeated acid phosphatase genes in yeast
 A:Reference number: S05794; MUID:85037940
 A:Accession: S05795
 A:Molecule type: DNA
 A:Residues: 1-467 <BAJ>
 A:Cross-references: EMBL:X01079; NID:94162; PIDN:CAA25555.1; PID:g758282
 A:Note: the authors translated the codon TAC for residue 272 as Thr
 A:Accession: A38792
 A:Molecule type: protein
 A:Residues: 18-45 <BAJ2>
 R:Mannhaupt, G.; Stucka, R.; Ehnhle, S.; Vetter, I.; Feldmann, H.
 Yeast 10, 1363-1381, 1994
 A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
 A:Reference number: S48255; MUID:95208357
 A:Accession: S48260
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-467 <FE2>
 A:Cross-references: EMBL:X78993; NID:9476045; PIDN:CAA55598.1; PID:g476051
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
 R:Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45927
 A:Accession: S45961
 A:Molecule type: DNA
 A:Residues: 1-467 <FE2>
 A:Cross-references: EMBL:X78993; NID:9476045; PIDN:CAA55598.1; PID:g476051
 R:Arima, K.; Oshima, T.; Kubota, I.; Nakamura, N.; Mizunaga, T.; Toh-e, A.
 Nucleic Acids Res. 11, 1657-1672, 1983
 A:Title: The nucleotide sequence of the yeast PHO5 gene: a putative precursor of repress
 A:Reference number: A00777; MUID:83168913
 A:Accession: A00777
 A:Molecule type: DNA
 A:Residues: 1-35, 'Y', 37-129, 'G', 131-293, 'O', 295-445, 'V', 447-461, 'DT', 464-465, 'K', 467 <AR
 A:Cross-references: EMBL:V01320; NID:94158; PIDN:CAA24630.1; PID:g4159
 A:Accession: A38793
 A:Molecule type: protein
 R:Meyhack, B.; Baizha, W.; Rudolph, H.; Hinnen, A.
 EMBO J. 1, 675-680, 1982
 A:Title: Two yeast acid phosphatase structural genes are the result of a tandem duplicat
 A:Reference number: S41855; MUID:84236032
 A:Accession: S41855
 A:Molecule type: DNA
 A:Residues: 1-30, 'T', 32-51, 'S', 53-75 <MEY>
 A:Cross-references: EMBL:M24178; NID:9172156; PIDN:AAA34868.1; PID:g172157
 R:Rait-Kamradt, A.G.; Turner, K.J.; Kramer, R.A.; Elliott, Q.D.; Bostian, S.J.; Thill, G
 Mol. Cell. Biol. 6, 1855-1865, 1986
 A:Title: Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene cluster within
 A:Reference number: A93074; MUID:87064474
 A:Accession: B25241
 A:Molecule type: DNA
 A:Residues: 1-44 <TAI>
 R:Berghman, L.W.
 Mol. Cell. Biol. 6, 2298-2304, 1986
 A:Title: A DNA fragment containing the upstream activator sequence determines nucleosom
 A:Reference number: A25367; MUID:87064526
 A:Accession: A25367
 A:Molecule type: DNA
 A:Residues: 1-2, 'Y', 4-43, 'T', 45-51 <BER>
 R:Silve, S.; Monod, M.; Hinnen, A.; Haguenauer-Tsapis, R.
 Mol. Cell. Biol. 7, 3306-3314, 1987
 A:Title: The yeast acid phosphatase can enter the secretory pathway without its N-termin
 A:Reference number: A27774; MUID:88036886
 A:Accession: A27774
 A:Molecule type: DNA

A:Residues: 1-51, 'S', 53-60 <SIL>
 A:Cross-references: GB:M17306
 C:Genetics:
 A:Gene: SGD:PHO5; MIPS:YBR093c
 A:Cross-references: SGD:S0000297; MIPS:YBR093c
 A:Map position: 2R
 A:Note: YBR093c
 C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-467/Product: acid phosphatase, repressible #status experimental <MAT>
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,103,162,192,250,315,356,390,439,445,456,461/Binding site: carbohydrate (Asn) (co
 F:337/Active site: His #status predicted
 Query Match 15.9%; Score 392; DB 1; Length 467;
 Best Local Similarity 25.6%; Pred. No. 9.8e-24;
 Matches 123; Conservative 64; Mismatches 216; Indels 78; Gaps 15;
 QY 6 VLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCPEISHLWGTVPYFSLADESAISP 65
 Db 5 VVYSILAASLANAGTI--PLGKLADVDTKIGTKDIPFPL---GGAGPYSPFGDYGISR 58
 QY 66 DVPDCRVTQVQLSRHGRYPTSSASKAYSALIEAIQNAFAKGYAFK-TYNTLIG 124
 Db 59 DLPEGCEMKOLOMVGRRGERTVSLAKTIKSTWYKLSNTYTRQFNGSLSLNDYEFFIR 118
 QY 125 ADO-----LTPF-GENQMVNSGKIFRYRYKALARKIVPF-IRASGSDRVIA 168
 Db 119 DDDLEMETTFANSDVLPNPTGEMNAKRHARDFLAQYGYMVENQTSFAVFTSNKRCMD 178
 QY 169 SAKFTIEGFSQAKLDPGQHPQASPVINVII-----PEGSGYNNTLDHGTCTCTAFEDSEL 223
 Db 179 TAQYFDGL-----GQD-----FNITQTVSEASAGANTLSACNSCPAW-DYDA 222
 QY 224 GDVEANFTALFAPAIRLEADLPGLTDEVDVYLMDCPFDVVARSDATELSPFCA 283
 Db 223 NDDIVNEYDTYLDLDDIAKRLNKENKGLNTSTDAFLFSWCAFEV-----NAGYSDDVCD 277
 QY 284 LFTHDEWIQDYDLSLQKGYGACAGNPLGPAQGVGFANELLARLTHSPVQDHTSTNHTLD 343
 Db 278 IFTKDELVLVHYSDYDHTYHGGPYDIKSGVSNLFNASVLLKQSEIQDQ----- 329
 QY 344 SNPATFPLNATLADFSHONTMISIFFALGLVNGTKPLSTTSVESIEETDGYSSASWTVPF 403
 Db 330 -----KWLSTFHTDIDILNLTAGIIDDKNLTAETVPPMGNT--FHRSWYVQ 377
 QY 404 AARAYVEMMQCAEKPLVKLVNDRVPLHGCAGVDKLGCRKRDFF-----VEGLSFA 456
 Db 378 GARVYTEKFCQ--SNDTYRVYVINDAVVPIETCTGTGFGFSCDEINDFYDAEKRVAGTDFL 435
 QY 457 R 457
 Db 436 K 436
 RESULT 7
 PABYCC
 N:acid phosphatase (EC 3.1.3.2) precursor, constitutive - yeast (*Saccharomyces cerevisi*
 N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 30-Sep-1991 #sequence revision 09-Sep-1994 #text change 05-Nov-1999
 C:Accession: S48259; S45960; S05794; A25241; S44674
 R:Mannhaupt, G.; Stucka, R.; Ehnhle, S.; Vetter, I.; Feldmann, H.
 Yeast 10, 1363-1381, 1994
 A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
 A:Reference number: S48255; MUID:95208357
 A:Accession: S48259
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-467 <MAN>
 A:Cross-references: EMBL:X78993; NID:9476045; PIDN:CAA55597.1; PID:g476050

A; Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R; Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
submitted to the Protein Sequence Database, August 1994
A; Reference number: S45927

```

RESULT      8
JC4285
N:Alternate names: orthophosphoric-monoester phosphohydrolase (acid optimum); Pho
C:Species: Pichia pastoris
C:Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999
C:Accession: JC4285
R:Payne, W.E.; Gannon, P.M.; Kaiser, C.A.
Gene 163, 19-26, 1995
A:Title: An inducible acid phosphatase from the yeast Pichia pastoris: Characterizati
A:Reference number: JC4285; MUID:96001238
A:Accession: JC4285
A:Molecule type: DNA
A:Residues: 1-468 <PAY>
A:Cross-references: GB:U28658; NID:g881955; PIDN:AAA85503.1; PID:g881956
A:Experimental source: GS115
C:Genetics:
A:Gene: phol
C:Superfamily: yeast acid phosphatase
C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydr
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-468/Product: acid phosphatase #status predicted <MAT>
F:84/Active site: His (phosphohistidine intermediate) #status predicted
F:163,196,256,321,360,453/Binding site: carbohydrate (Asn) (covalent) #status predict
F:345/Active site: His #status predicted

```

RESULT 9
S52495
acid phosphatase homolog YDI024c - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein D2815

A;Residues: 1-468 <URR>
A;Cross-references: EMBL:Z74072; NID:g1430996; PIDN:CAA98583.1; PID:g1430997; MIPS:YDL021
A;Experimental source: strain S288C
C;Genetics:
A;Map position: 4L
C;Superfamily: yeast acid phosphatase

Qy 108 AFKGKATLKTYNTL-----GADDLT-PF-GENQMVNSGIFFRYRYKAL--ARKIVPF 157
 :| | | : : : : : : : : : : : : : : :
Db 109 FYKGDIAFINDTWTVYVPNECYVNAETSGPVAGLIIDAYHNGNDYKARYCHIWNETGVDPF 168

Qy 158 IRASGSDRVIASAEKFFIEGFQSAKLADPGSQPHQASPVINVII-PEGSGYNNTLDHGTC**T** 216
 ::| | | | : | | | | : : | | | | :
Dd 169 F-SSGYGRVIETARKFGEGFECY-----YSTNAALNIISEVMGADSLTP--TCD 217

Qy	217	AFEDSELGDDVEANFTALFAPAIRLEADLPGVTLTDEDVVLMDCPFDTVARTSDAT	276
		<div style="display: flex; justify-content: space-around; width: 100%;"> </div>	
Db	218	TNDQTTCDNLTYQLPQFKVAA--ARLNSQNGMNLTASDVYNLMVMASFELNAR----	270

Qy	277	ELSPFCALFTHDEWIQDYDLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVDHT	336
Db	271	PFSNWINAFTQDEWVSFGYVEDLNYYYCAGPDKNMAAVGAVYANASLTLLNQGPKE---	327

QY 337 SINHTLSDNPATPLNATLYADESHDNTMISIFFALGLY--NGTKPLSTTSVSETEETDG 394
 Db 328 -----AGSLFFNFAHDNTNITPLAALGVLIPNEDLPL-----DRVAFGNP 367

QY 395 YSASWTVPFAARAYVEMMCOA----EKEPLVRVLVNDVRVPLHGC 436

Db 368 YSIGNIVPMGGHLTIERLSQATALSDEGTYVRLVLEAVLPFND C 413

RESULT 11
JN0715
3-phutase (EC 3.1.3.8) precursor - *Aspergillus fumigatus*

N;Alternate names: pH 2.5-optimum acid phosphatase
C;Species: *Aspergillus ficuum*
C;Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
C;Accession: JN0715, PN0594, PN0460

A:Reference number: JN0715; MUID:93371452

A;Accession: JN0713
A;Molecule type: DNA
A;Residues: 1-479 <EHR>
A;Cross-references: GB:L20567

A; Molecule type: protein
A; Residues: 20-101; 133-146; 376-399 <EH2>
R; Ullah, A.H.J.; Dischinger Jr., H.C.

A;Title: Identification of active-site residues in Aspergillus ficuum extr
A;Reference number: PN0460; MUID: 93249452
A;Accession: PN0460

A: Molecule type: protein
A: Residues: 65-66, 68-93 <ULL>
C: Comment: This enzyme, previously characterized as an acid phosphatase (EC 3.1.3.2), has
C: Genes: phbB
A: Gene: phbB
A: Introns: 261/1; 300/2; 335/2
C: Superfamily: yeast acid phosphatase
C: Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
F: 1-19/Domain: signal sequence #status predicted <SIG>
F: 20-479/Product: 3-phytase #status experimental <NAT>
F: 81,337/Active site: Arg, His #status predicted
F: 82/Active site: His (phosphohistidine intermediate) #status predicted
F: 106,191,227,250,315,340,425,442,458/Binding site: carbohydrate (Asn) (covalent) #status
Query Match 14.1%; Score 346.5; DB 1; Length 479;
Best Local Similarity 25.3%; Pred. No. 4.7e-20;
Matches 119; Conservative 54; Mismatches 170; Indels 63; Gaps 17;
Qy 49 GTSPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAY-SALTEAIQKNAT 107
Db 53 GPYSERVSY-----GIADPPPGCEVDQVIMVYKRGHGERVPSAGKSIEEALAKVYSINT 108
Qy 108 AFKGYAFKTYNYTL-----GADDLT-PF-GENQMVNSGKIFRYKAL--ARKIYVPF 157
Db 109 EYKGLAFPLNDWTYYVNPCEYNAETTSQVYAGLLDAYNHGNDYKARYGHLWNGETVVPF 168
Qy 158 IRASGRDVIASAEKFIQFOSAKLADPGSQPHQASPVINII-PEGSGYNTLDHGCT 216
Db 169 F-SSGVRVIETARKGEGFGYN-----YSTNAALNIISEVNGADSLP--TCD 217
Qy 217 AFESELDGDDVEANFTALFAIPARLEADLPVLTDEDVYVLMDCPFDFTVARTSDAT 276
Db 218 TDNDTTCNLTYQLPQFKVAA--ARLNSQPNGLTASDVYVLMVMAFELNAR----- 270
Qy 277 ELSPPCALFTDHWITQYDYLQSLGKYGYGAGNPLGPAQGVGFANELLARLTHSPVDHT 336
Db 271 PFSNMINAFTQDQWVSFGYEDLNYYCAGPGDKNMAAGVAVYANASLTLLNQGP----- 325
Qy 337 STNHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLY--NGTKPLSTTSVESTETD 394
Db 326 -----KEAGP-----LFTLAHDNTITPILAAGLVILPNEIDLPL-----DRVAFGNP 367
Qy 395 YSASWTVPFAARAYVEMMQCOA-----EKEPLRVRLVNDRVVPLHGC 436
Db 368 YSIGNIVPWGHLTIERSQCATLSQATGTYVRLVINEAVLPFND 413
RESULT 12
T39929
thiamin-repressible acid phosphatase - fission yeast (Schizosaccharomyces pombe)
C: Species: Schizosaccharomyces pombe
C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C: Accession: T39929
R: Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Xiang, Z.; Aves, S.
submitted to the EMBL Data Library, May 1998
A: Reference number: 221857
A: Accession: T39929
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-463 <LYN>
A: Cross-references: EMBL:AL023286; PIDN:CAA18863.1; GSPDB:GN00067; SPDB:SPBC21H7.03c
A: Experimental source: strain 972h; cosmid c21H7
C: Genes: phbB
A: Gene: SPDB:SPBC21H7.03c
A: Map position: 2
C: Superfamily: yeast acid phosphatase

Qy 46 HLWGTYS-----PYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTS--SASKAYSA-- 97
Db 35 HL-GTSPYHHPYFGLDSA-----FPECEIQVHLLQRHGSRNPTGDVTATDVYSSQY 88
Qy 98 -----LIEAIQKNATAFKGYAFKTYNYTL-----GADDLTTPFGENQMVNSGKIFRYR 146
Db 89 LNNFOEKLNGSIPVNFSPENPLCFIKQWTPVIDAENADQLSSRGRLFLDLGRQLYOR 148
Qy 147 YKALARKIVPFRASGRDVIASAEKFIQFOSAKLADPGSQPHQASPVINIIPEG--S 204
Db 149 YKLFDSVYDINTAEQERVSESAAKFTYGLFGDMYKTN-----FILISEGKAA 199
Qy 205 GYNNTLDHGTCTAFDSEL-----GDVDEANFTALFAIPARLEADLPVLTDEDVYV 259
Db 200 GANSLSMYNACPVFKDNFHNKATDAHAVRNIEFIVNRLAKYFSSYKLTINDVRS 259
Qy 260 LMDMCPFDFTVARTSDATLSPPCALFTDHWITQYDYLQSLGKYGYGAGNPLGPAQGVGF 319
Db 260 LFYICEYEYATK-----DHSDFCSITPSEFLNFEYDSDLDQAYGGPVSEWASTLGGAY 314
Qy 320 ANELLARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLYNGTK 379
Db 315 INNADSL-----RNVNTPDFDRK-----VFLAFTHDSNIIPEVAALGFPPDIT 358
Qy 380 PLSTTSVESIEETDGYASWTVPFAARAYVEMMQCOAEPVLRVRLVNDRVVPLHGC 439
Db 359 PONPLTPDKNIYTSQKTSFVFPFAGNLITELFFC-SDSKYVRLVHVNQVYVPLDCCGYG 417
Qy 440 KLGR-----CKRDFVFEGLSFARSGGNWAECE 466
Db 418 PSGTSDGLCELQAYLNSPIRANSTNGISIF 448
RESULT 13
A25326
acid phosphatase (EC 3.1.3.2) precursor [similarity] - fission yeast (Schizosaccharom
C: Species: Schizosaccharomyces pombe
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C: Accession: A25326; T50405
R: Elliott, S.; Chang, C.; Schweingruber, M.E.; Schaller, J.; Rickli, E.E.; Carbon, J.
J. Biol. Chem. 261, 2936-2941, 1986
A: Title: Isolation and characterization of the structural gene for secreted acid phos
A: Reference number: A25326; MUID:86140050
A: Accession: A25326
A: Molecule type: DNA
A: Residues: 1-453 <ELL>
A: Cross-references: GB:M11857; NID:g173422; PIDN:AAA35321.1; PID:g173423
R: Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 2000
A: Reference number: 225068
A: Accession: T50405
A: Molecule type: DNA
A: Residues: 1-453 <RIE>
A: Cross-references: EMBL:AL137099; PIDN:CA86857.1; GSPDB:GN00067; SPDB:SPBP4G3.02
A: Experimental source: strain 972h(-); clone p1 p4G3
C: Genes: phbB
A: Gene: phbB; SPDB:SPBP4G3.02
A: Map position: 2
C: Superfamily: yeast acid phosphatase
C: Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F: 68/Active site: Arg #status predicted
F: 69/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 12.1%; Score 297; DB 1; Length 453;
Best Local Similarity 24.8%; Pred. No. 4.3e-16;
Matches 99; Conservative 60; Mismatches 190; Indels 50; Gaps 9;
Qy 51 YSPYFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSA-----LIE 100
Db 43 HKPYF-----YGPSIDFPTTKIKQVHTLQRHGSRNPT--GGNAAFDAVGIANFQORLLNG 96
Qy 101 AIQKNATAFKGYAFKTYNYTL-----GADDLTTPFGENQMVNSGKIFRYRKYKALARKIVP 157

Db 97 SVPIIDYSGNPLSFVPTWTPVIEANADALSSGRVLEFDMGRQFYERYHELFENASTYN 156
QY 158 IRASGSDRVIAAEKFIIEGQSAKLADPGSQHQPSPVINVIIPGSGYNNITLDHGTCTA 217
Db 157 IYTAQQRVDSALMYGGMF-----GEDVHNFNTYILVSENATAGSNLSYNAACP 209
QY 218 FEDSELGDDVEANFTALPAPAIRARLEADLPVTLTDEDVVVLMDCPDDTVARTSDATE 277
Db 210 SDADEFTTALCAWRNVYPPRQRLNFFSNVNTINDILNYGICSYEIALQ-----D 264
QY 278 LSPFCALFTHDEWIOYDIQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTS 337
Db 265 YSEFCKLNSVDFLNFYEGDLFSYGMGNSVKWGSIFGGAYANSLSL--RSVENNTQ 322
QY 338 TWHITLDSNPATPLNATLYADFSHNTMISIFFAALGLYNGTGPLSTTSVESIEETDGYSA 397
Db 323 -----YQVFFAFTHDANIIPVETALGFFTDNTPENPLETSYQVHSHMKA 366
QY 398 SWTPVFAARAYVEMMOCAKEPLRVLRVLRVPLHGC 436
Db 367 SEFVFAGNLITELFQCEDSKI-YVRHLVNEVFPPLSDC 404

RESULT 14
S14119
acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
C:Accession: S14119; T40455
R:Yang, J.; Schweingruber, M.E.
Curr. Genet. 18, 269-272, 1990
A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schizosaccharomyces pombe
A:Reference number: S14119; MUID:91064763
A:Accession: S14119
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-463 <YAN>
A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA40258.1; PID:g5007
R:Lyne, M.; Rajandream, M.A.; Barrall, B.G.; Baker, S.; Mungall, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21931
A:Accession: T40455
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-463 <LYN>
A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN000067; SPDB:SPBC428.03c
A:Experimental source: strain 972h-; cosmid c428
C:Genetics:
A:Gene: SPBC428.03c
A:Map position: 2
C:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 11.3%; Score 278.5; DB 2; Length 463;
Best Local Similarity 25.5%; Pred. No. 1.4e-14;
Matches 108; Conservative 65; Mismatches 187; Indels 63; Gaps 15;
QY 46 HLWGYG----PYFLADESAISPDVDCRVTFVQVLSRHGARYPT--SSASKAYSA-- 97
Db 35 HL-GTISVYHEPF-----NGPTTSFPESCAIKQVHLLQRHGSRNPTGDDTATDVSSAQ 88
QY 98 -----LIFAIQKNATAFKGYAFLKTYNTL---GADDITPGENQMVNSGKIFKFR 146
Db 89 IDIFONKLLNGISYFNSTPENPLFYVHWTPVIAKNAQDLSGRIELFDLGRQVFER 148
QY 147 YKALARKIVPFIIRASGSDRVIAAEKFIIEGQSAKLADPGSQHQPSPVINVIIPGSGYNNITLDHGTCTA 217
Db 149 YVELFDTDYDINTAAQERVVDSAEWFSYGMFGDDMQNKTN-----FIVLPEDDSA 199
QY 205 GYNNTLDHGTCTAFDESELGDD--VEANFTA----LFAPAIRARLEADLPVTLTDEDVYV 259

Db 200 GANSLAMYSCPVYEDNNIDENITEAAHTSWRNVLKPIANRLNKYFDSGYNLTVSDVRS 259
QY 260 LMDMCPFDVTARTSDATLSPFCALFTHDEWIOYDIQSLGKYGYGAGNPLGPAQGVGF 319
Db 260 LYVICVYIEALRDN-----SDFCSLFTPSEFLNFYDSDLDYAYWGGPASEWASTLGGAY 314
QY 320 ANELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMISIFFAALGLYNGTK 379
Db 315 VNNLANNL-----RKGVNNASDRK-----VFLAFTHDSQIIPVEAALGFEPDIT 358
QY 380 PLSTTSVESIEFTDGYSAISWTVPFAARAYVEMMOCAKEPLRVLRVLRVPLHGCAYD 439
Db 359 PBHPLPTDKNIFTYSLKTSFVFPFAGNLITELFLC--SDNKYYVVRHLVNCQVYPLTDCGYG 417
QY 440 KLG 442
Db 418 PSG 420

RESULT 15
JE0369
histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JE0369
R:Mullaney, E.J.; Ullah, A.H.J.
Biochem. Biophys. Res. Commun. 251, 252-255, 1998
A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidopsis
A:Reference number: JE0369; MUID:99009256
A:Accession: JE0369
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <MUL>
C:Superfamily: yeast acid phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 6.1%; Score 149; DB 2; Length 465;
Best Local Similarity 22.3%; Pred. No. 0.00037;
Matches 101; Conservative 63; Mismatches 199; Indels 90; Gaps 22;
QY 35 DGGYQCFPEISHLMGTYSPIYFSLAD--ESAI-SPDVPDDCRVTFVQVLSRHGARYPTSSA 91
Db 16 DGGF----DVRHHLSTVTRYSTSKDVTQNLIEGSNVPSSECTPIHLNLVARHCTRSPTKKR 71
QY 92 SKAYSAL---TEAIQKNATAFK-----GKYAFLKTYNTYTLGADLLTPFGENQMVN 138
Db 72 LREMESLAGRFKELVRDAEARKLPDKIPGWLQGWK--SPWEGKVKGGLIRQGEDELYQ 129
QY 139 SGIKFYRRYKALARK---IVPFIASGSDRVIAAEKFIIEGQSAKLADPGSQHQPAS 194
Db 130 LGIRVRERFPLFEDYHPDVYTIIRATQIPRASASAVAFGMGLFSEK-----GNL----- 179
QY 195 VINVIIPGSGYNNITLDHGTCTAFDESELGDDVEANFTALPAPAIRARLEADLPVTLT 254
Db 180 -----GPGNRNFA-----AVTSENKASDTKLRFEC--CONYKSKYKAKEPAVDK 224
QY 255 EDVYVLMDC-----PDDTVARTSDATELSPPFCALFTHDEWIOYDIQSLGKYGYGAGNPL 311
Db 225 EPLANKITASVAKRYDLKFTKODISSLMFLCKQVALLW-----TDDLEVLFLKGYGNSL 279
QY 312 GPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMISIFFA 371
Db 280 NYKMGV-----PLLEDVLHSMEEAIKAREKLP--PGSYE-KARL--RFAHAETIIVPPFSC 330
QY 372 -LGLYNGTKPLSTTSVESIE-----ETDGYSAISWTVPFAARAYVEMMOCAKE 418
Db 331 LIGLF-----LDGSEFEKIQREKPLELPQPPKTRDRFGSTMAPFGGNNILVLYSCPAES 385
QY 419 EP--LVRLVLRVLRVPLHGCAYDVKLGRCKRDF 449
Db 386 SPKYEQVQLHNEHPITAVPGC--DGKDFCPLDEF 416

Search completed: October 26, 2001, 15:10:29
Job time: 6335 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:18:03 ; Search time 29.88 seconds
(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGFFVLLSTATLFGTSET.....DFVEGLSFARSGNNAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1879	76.3	467	1	PHYA_ASPAW	P34753 aspergillus
2	1875	76.2	467	1	PHYA_ASPNG	P34752 aspergillus
3	1835	74.6	463	1	PHYB_EMENI	O00093 emericeilla
4	395	16.0	467	1	PPAB_YEAST	P35842 saccharomyc
5	395	16.0	467	1	PPAC_YEAST	P38693 saccharomyc
6	392	15.9	467	1	PPAS_YEAST	P00635 saccharomyc
7	380	15.4	467	1	PPA3_YEAST	P24031 saccharomyc
8	355	14.4	468	1	PPAL_PICPA	P52291 pichia past
9	354	14.4	468	1	PPAD_YEAST	P52290 saccharomyc
10	348.5	14.2	479	1	PHYB_ASPNG	P34754 aspergillus
11	347.5	14.1	479	1	PHYB_ASPAW	P34755 aspergillus
12	341.5	13.9	469	1	PPA5_KLULA	P52289 kluyveromyc
13	297	12.1	453	1	PPA1_SCHPO	P08091 schizosacch
14	278.5	11.3	463	1	PPA2_SCHPO	Q01682 schizosacch
15	130	5.3	735	1	PPAX_CAEEL	Q09549 caenorhabdi
16	124.5	5.1	421	1	PPAL_MOUSE	P24638 mus musculu
17	114.5	4.7	423	1	PPAL_RAT	P20611 rattus norv
18	113	4.6	874	1	COPG_HUMAN	Q9v678 hmo sapien
19	112	4.5	721	1	NICA_CAEEL	Q23316 caenorhabdi
20	102.5	4.2	423	1	PPAL_HUMAN	P11117 homo sapien
21	101.5	4.1	680	1	NCPR_CANMA	P50126 candida mal
22	101	4.1	874	1	COPG_BOVIN	P53620 bos taurus
23	98.5	4.0	381	1	PPAP_RAT	P20646 rattus norv
24	96.5	3.9	275	1	UL11_HCMVA	P16721 homo cypien
25	96.5	3.9	386	1	PPAP_HUMAN	P15309 hmo sapien
26	96	3.9	411	1	PPAY_CAEEL	Q10944 caenorhabdi
27	96	3.9	413	1	PPAW_CAEEL	Q09451 caenorhabdi
28	95.5	3.9	680	1	NCPR_CANTR	P37201 candida tro
29	95	3.9	605	1	AMY_THECU	P29750 thermomonos
30	94.5	3.8	580	1	UL25_HSV11	P10209 herpes simp
31	94.5	3.8	602	1	PEHX_ERWCH	P15922 erwinia chr
32	94.5	3.8	760	1	MCMT_SCHPO	Q75001 schizosacch
33	94	3.8	754	1	MBPI_KLULA	P39679 kluyveromyc

ALIGNMENTS

RESULT 1

ID	PHYA_ASPAW	STANDARD;	PRT;	467 AA.
AC	P34753;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE			
DE	3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE			
DE	PHOSPHOHYDROLASE A).			
GN	PHYA OR PHY.			
OS	Aspergillus awamori.			
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.			
OX	NCBI_TaxID=105351;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ALK0243;			
RX	MEDLINE=94040796; PubMed=8224894;			
RA	Piddington C.S., Houston C.S., Palohimo M.T., Cantrell M.A.,			
RA	Miettinen-Oinonen A., Nevalainen H., Rambosk J.A.;			
RT	"The cloning and sequencing of the genes encoding phytase (phy) and			
RT	pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.			
awamori."				
RL	Gene 133:55-62(1993).			
CC	-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE			
CC	FROM PHYTASE.			
CC	-!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =			
CC	D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.			
CC	-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; L02421; AAA16898.1; -			
DR	PIR; JN0889; JN0889.			
DR	HSSP; P34752; LIHP.			
DR	InterPro; IPR000560; -			
DR	Pfam; PF00328; acid_phosphat; 1.			
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.			
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.			
DR	Hydrolase; Glycoprotein; Signal.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	467	3-PHYTASE A.
FT	ACT_SITE	82	82	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE	361	361	PROTON DONOR (BY SIMILARITY).
FT	DISULFID	31	40	BY SIMILARITY.
FT	DISULFID	71	414	BY SIMILARITY.
FT	DISULFID	215	465	BY SIMILARITY.
FT	DISULFID	264	282	BY SIMILARITY.
FT	DISULFID	436	444	BY SIMILARITY.

p30002 human herpe
p52340 human herpe
Q52309 providencia
O13712 schizosacch
O42662 solenostemo
P19584 thermoanaer
Q08193 saccharomyc
P09870 clostridium
Q10061 schizosacch.
Q69014 manchester
P38928 saccharomyc
O04931 beta vulgar

```

FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 51075 MW; 118E828A5D7EC661 CRC64;

Query Match 76.3%; Score 1879; DB 1; Length 467;
Best Local Similarity 75.68; Pred. No. 6.3e-143;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGCVFVLLSIATLFGSTGTALPRGNHSDCTVDGGYQCFPEISHLWGTVPYFSLADE 60
Db 1 MGVSALLPLVLLAGVTSGLVAPASRNQSTCDTVDQGYQCFSETSHLWGVAPFFSLANE 60

QY 61 SAISPDVDDCRVTQVLSRHGARYPTSSAKAYSLIEIAIQNATAFKGYAFLKTYN 120
Db 61 SAISPDVAGCRVTTPAQLSRHGARYPTESKGYKYSALIEIQNVTTFDGKYAFLKTYN 120

QY 121 YTLGADLTTPFGENOMVNSGKIFVRRYKALARKIVPFRASGSDRVIASAEKFTGFSQA 180
Db 121 YSLGADLTTPFGEGLVNSGKIFVRYESLIRNIPFIRSSGSRVSIASGEKFTGFSQT 180

QY 181 KLADPGSQPHQASPVINIIPEGSGYNNLTDLHGCTAFEDSELGDDVDVEANFTALFAPAIR 240
Db 181 KLKDPRAQPGQSSPKIDVIVSEASSNNTLPGTCTFEDSELADTVEANFTATFAPSIR 240

QY 241 ALEADLPGVTLTDEBVLMDMCPDFVARTSDATELSPECLFTHDEWTOYQYDLSLG 300
Db 241 QRLNDLSGVTLTDEVTYLMDCSFDISTSTVDTKLSPCDFLTHDEWTHYDLSLK 300

QY 301 KYGCGAGNPLGPAQGVANELIARLHSPVODHTSNHPLNSPATFPLNATLYADFES 360
Db 301 KYGCGAGNPLGPTQGVANELIARLHSPVHDTSNNHPLNSPATFPLNATLYADFES 360

QY 361 HDNTMISFALLGNGKPLSTTSVESIEETDGYSAWSTVPPFAARAYVEMMQCAKEP 420
Db 361 HDNGIISILFALGNGKPLSTTVENITQDGFSSAWTVPFASRLYVEMMQCAQOE 420

QY 421 LVRVLNDRVPLHCAVDKLGCRKRDFFVGLSFARSGGNWAECPA 467
Db 421 LVRVLNDRVPLHGPIDALGRCTRDSEFVRGLSFARSGGDWAECSA 467

RESULT 2
PHYA.ASPNG
ID PHYA.ASPNG STANDARD; PRT; 467 AA.
AC P34752;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.6) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUM;
RX MEDLINE=93252284; PubMed=8387447;
RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
RA Gouka R.J., Suykerbuyk M.E.G., Luiten R.G.M., van Paridon P.A.,
RA Sellen G.C.M., Veenstra A.E., van Gorcom R.F.M.,
RA van den Hondel C.A.M.J.J.;

```

```

RT "Cloning, characterization and overexpression of the phytase-encoding
RT gene (phyA) of Aspergillus niger.";
RL Gene 127:87-94(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mullaney E.J.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-464.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUM;
RX MEDLINE=93249451; PubMed=8387289;
RA Ullah A.H.J.; Dischinger H.C. Jr.;
RT "Aspergillus ficum phytase: complete primary structure elucidation
RT by chemical sequencing.";
RL Biochem. Biophys. Res. Commun. 192:747-753(1993).
RN [4]
RP SEQUENCE OF 71-93.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUM;
RX MEDLINE=91298982; PubMed=1648914;
RA Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
RT "Cyclohexanediol modification of arginine at the active site of
RT Aspergillus ficum phytase.";
RL Biochem. Biophys. Res. Commun. 178:45-53(1991).
RN [5]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUM;
RX MEDLINE=89160685; PubMed=2852807;
RA Ullah A.H.J.;
RT "Aspergillus ficum phytase: partial primary structure, substrate
RT selectivity, and kinetic characterization.";
RL Prep. Biochem. 18:459-471(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUM;
RX MEDLINE=97307250; PubMed=9164457;
RA Kostrewa D., Gruninger-Leitch F., D'Arcy A., Broger C., Mitchell D.,
RA van Loon A.P.;
RT "Crystal structure of phytase from Aspergillus ficum at 2.5-A
RT resolution.";
RL Nat. Struct. Biol. 4:185-190(1997).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -1- BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN
CC FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
CC USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER
CC THE NAME PHYTASE NOVO.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z16144; CAA78904.1; -.
DR EMBL; M94550; AAA32705.1; -.
DR PIR; JN0482; JN0482.
DR PIR; JN0023; JN0023.
DR PIR; JN0656; JN0656.
DR PDB; 1IHP; 18-MAR-98.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 467 3-PHYTASE A.
FT ACT_SITE 82 82 NUCLCOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).

```

FT DISULFID 31 40
 FT DISULFID 71 414
 FT DISULFID 215 465
 FT DISULFID 264 282
 FT DISULFID 436 444
 FT CARBOHYD 27 27
 FT CARBOHYD 59 59
 FT CARBOHYD 105 105
 FT CARBOHYD 120 120
 FT CARBOHYD 207 207
 FT CARBOHYD 230 230
 FT CARBOHYD 339 339
 FT CARBOHYD 352 352
 FT CARBOHYD 376 376
 FT CARBOHYD 388 388
 FT SEQUENCE 467 AA; 51086 MW; 88FE8F3584341D6D CRC64;

Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 1.3e-142;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTLGPRGNHSDCTVDGGYQCFPEISHLWGTSPYFSLADE 60
 Db 1 MGVSAVLLPLVLLSGVTSLGLAVPASRNGSCDVTVDGGYQCFSESHLWGYAPFSLANE 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
 Db 61 SVISPEVAGCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
 QY 121 YTLGADDTLTPGEGNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
 Db 121 YSLGADDTLTPGEGNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
 QY 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTAFDESELGDDVEANTALFAPAIR 240
 Db 181 KLKPRAQPGQSSPKIDWIVSEASSNNTLDHGCTAFDESELGDDVEANTALFAPAIR 240
 QY 241 ARLEADLPVGLTDEVDVYLMDMCPDFTVARTSDATLSFPCALFTHDEWIQDYLOSIG 300
 Db 241 QRLNDLSGVTLTDEVTYLMDCSFDTITSTVDTKSPFCDLFTHDEWIQDYLOSIG 300
 QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYYGAGNPLGPTQGVGYANELIARLTHSPVHDDTSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMQCAKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTSVENITQTDGFSAMTVFPASRLYVEMMQCAKEP 420
 QY 421 LVRLVNDRVVPLHGCADVRLGKCRKDDFVEGLSFARSGGNWAECEFA 467
 Db 421 LVRLVNDRVVPLHGCADVRLGKCRKDDFVEGLSFARSGGNWAECEFA 467

RESULT 3
 PHYB EMENI
 ID PHYB EMENI STANDARD; PRT; 463 AA.
 AC O00093;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE B)
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98007872; PubMed=9349716;
 RA Pasamontes L., Halker M., Henriquez-Huecas M., Mitchell D.B.,

van Loon A.P.;
 "Cloning of the phytases from Emericella nidulans and the
 thermophilic fungus Talaromyces thermophilus";
 Biochim. Biophys. Acta 1353:217-223(1997).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 FROM PHYTATE.
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H₂O =
 D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U59803; AAB96871.1; ..
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid_phosphat_1;
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 463 3-PHYTASE B.
 FT ACT_SITE 80 80 REQUIRED FOR BINDING SUBSTRATE
 FT ACT_SITE 81 81 (BY SIMILARITY).
 FT ACT_SITE 357 357 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 463 AA; 51786 MW; ECC5827D1E1C82A2 CRC64;

Query Match 74.6%; Score 1835.5; DB 1; Length 463;
 Best Local Similarity 74.0%; Pred. No. 1.9e-139;
 Matches 345; Conservative 46; Mismatches 70; Indels 5; Gaps 3;

QY 1 MGVFVLLSIATLFGSTGTLGPRGNHSDCTVDGGYQCFPEISHLWGTSPYFSLADE 60
 Db 1 MAFFTVALLSYLLSRVSAQA--PVVQNHSCNTADGGYQCFPNVSHVWGQYSPYFSEIE 58
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
 Db 59 SAISEDVPHGCEVTFVQVLSRHGARYPTESKAYSGLIEAIQKNATSFWGQYAFLESYN 118
 QY 121 YTLGADDTLTPGEGNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 180
 Db 119 YTLGADDTLTPGEGNSGKIFRYRKALARKIVPTIRASGSDRVIASAEKFEQFSA 178
 QY 181 KLADPGSOPHOASPVINVIPEGGSYNNTLDHGCTAFDESELGDDVEANTALFAPAIR 240
 Db 179 QLHDGSK--RAIPVNVNIIPEIDGFNNLTDHSTCVSFENDERADEIANFTMGPPIR 236
 QY 241 ARLEADLPVGLTDEVDVYLMDMCPDFTVARTSDATLSFPCALFTHDEWIQDYLOSIG 300
 Db 237 KRLENDLPGLITNENVYLMDCSFDTMARTAGTLELSPCALFTTEKEMQDYLOSLS 296
 QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 297 KYYGAGSPLGPAQGGIGFTNELIARLTQSPVQDNTSTNHTLDSNPATFPLDRKLYADES 356
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMQCAKEP 420
 Db 357 HDNSMISIFFAMGLYNGTQPLSDMSVESIQEMDGYAASWTVFPGARAYFELMQCE-KKEP 415

FT	CONFLICT	17	17	A -> L (IN REF. 2).
FT	CONFLICT	82	83	VS -> AR (IN REF. 2).
FT	CONFLICT	150	150	R -> H (IN REF. 2).
FT	CONFLICT	354	354	K -> Q (IN REF. 2).
FT	CONFLICT	423	423	D -> G (IN REF. 2).
SO	SEQUENCE	467 AA;	52757 MW;	AECDIC046B326G3 CRC64;

Query Match	16.0%;	Score 395;	DB 1;	Length 467;
Best Local Similarity	26.0%;	Pred. No. 3.7e-24;		
Matches 119;	Conservative 66;	Mismatches 207;	Indels 66;	Gaps

QY	24	PRGNHSCTVDGGYQCPEPEIHLHGTTSPFSLADESAISDPDCCDVRTVFQVLSRBHG	83
DB	21	PLGKLSIDIKIGTQIEIPFL----GGSPYYSPFEDGYISRDLEPCEMKQVOMVGRHG	76
QY	84	ARYPTSSASKAYSALEIAQKNATAFKGYAFLK-----TNYNLT-----GADDL	128
DB	77	ERYPTVSKAKSINTTWYKUSNTGQFSGALSFLNDDYEFFTRDTKNLEWETLANSVNVL	136
QY	129	TTF-GENQMVNSGIFRYRYKALARKIVPF-IRASGSDRVIASAEKFIEGFOSAKLADPG	186
DB	137	NPYTGENNAKRHARDFLAQGYGMVENQTSFVFTSNSNRCHDTAQVFDG-----LGDKF	191
QY	187	SQPHQASPVNIIIPGSGYNNLTDHGTCATFEDSELGDDVVEANTALFAPAIRARLEAD	246
DB	192	NISLQ-----TISEASAGANTLSAHHSCPAWDD-DVNDILKLYDTRKYSLGIAKRLNKE	245
QY	247	LPQVTLTDEDVYVLMDCPFDVARTSDATELSPPFCALETFHDEWIQYDYLQSLGKYGVYG	306
DB	246	NGELNLTSSDANTFFAWCAYENAR-----GYSDICNIETKDELVRFSYGQDLETYYQTG	300
QY	307	AGNPLGPAQGVGFANELIARLTHSPVODJHSTNNHTLDSNPATFPLNATLIYADFSDHNIMI	366
DB	301	PGYDVVRVSGANLFNASVKLLKESEYQDQ-----KVMLSFTHTDITDL	342
QY	367	SIFALGLYNGTRPLTTSVESTETDGYSSASWTVPFAARVYEMMQCAEPLRVIV	426
DB	343	NYLTIGIIDDKNLTAEHVPFFMENT--FHRSWVYVPGQARVYTEKFOC--SNDTYRVYVI	398
QY	427	NDRVVPJLHGCAVDKLGCRKDDPF-----VEGLSPAR	457
DB	399	NDAAVPTFTVCTGPGGSECLNDYEDYAEKRVAGTDFELK	436

RESULT	5			
PPAC_YEAST				
ID	PPAC_YEAST	STANDARD;	PRT;	467 AA.
AC	P38693;			
DT	01-FEB-1995	(Rel. 31, Created)		
DT	01-FEB-1995	(Rel. 31, Last sequence update)		
DT	01-FEB-1995	(Rel. 31, Last annotation update)		
DE	ACID PHOSPHATASE PHO12 PRECURSOR	(EC 3.1.3.2).		
GN	PHO12 OR YHR215W.			
OS	Saccharomyces cerevisiae	(Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina;	Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae;	Saccharomycetes.		
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N A.			
RC	STRAIN-S288C / AB972;			
RX	MEDLINE=94378003; Pubmed=8091229;			
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,			
RA	Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,			
RA	Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,			
RA	Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,			
RA	Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,			
RA	Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,			
RA	Vaudin M.;			
RT	"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome			
RT	VIII";			
RL	Science 265:2077-2082(1994).			
CC	-!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN			

CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- INDUCTION: S.CEREVISIAE HAS 2 TYPES OF ACID PHOSPHATASE: ONE IS
CC CONSTITUTIVE AND THE OTHER IS REPRESSIBLE BY INORGANIC PHOSPHATE.
CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00029; AB69729.1; -
DR PIR; S48996; S48996.
DR HSP; P34752; 1HP.
DR SGD; S0001258; PHO12.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat_1;
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 17 BY SIMILARITY
FT CHAIN 18 467 ACID PHOSPHATASE PHO12.
FT ACT_SITE 75 75 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 337 337 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 162 162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 467 AA; 52699 MW; BEC606CDF39B845B CRC64;

Query Match 16.08; Score 395; DB 1; Length 467;
Best Local Similarity 26.08; Pred. No. 3.7e-24;
Matches 119; Conservative 66; Mismatches 207; Indels 66; Gaps 13;
QY 24 PRGSHSCDTVDGQYCFPELISHLWGYSPYFSLADSAISDPDPCRTVFQVLSRHG 83
Db 21 PLGLSDIDKIGTQTEIFPFL---GGSGPYSPGPGYGISRLPESCEMKQVMVGRHG 76
QY 84 ARYPTSSASKAYSALIEAIOKNATAFKGVAFK-----TYNYTL-----GADDL 128
Db 77 ERYPTVSKAKSIMTWKLSNYTQFSGALSFLNDYEFTRDTKLEMETTLANSVNVL 136
QY 129 TPF-GENOMVNSGKIFRYRYKALARKIVPF-IRASGSDRVIAAEKFIQFQSAKLADPG 186
Db 137 NPYTGEMNAKRHARDFLAQYGMVYENQTSFAVTSNSNRCHDTAQYFDG-----LGDKF 191
QY 187 SQPQASPVNVIPEGSYNTLDHCTCFADSELGDDVEANFTALFAPARLEAD 246
Db 192 NISLQ-----TISAESAGANTLSAHSRCPAWDD-DVNDILKKYDKTKYLSGIAKRLNKE 245
QY 247 LPQVTLTDEDVYVLMDCPFDVTARTSDATLSPFCALFTHDEWIDYQDYLSQKGYGVG 306
Db 246 NKGLNLTSSDANTFAWCAEINAR-----GYSDICNIFTKDELVRFSYQDLETYYQTG 300
QY 307 AGNPLGPAQGVGFANELIARLTSPVDHPTSTNHTLDSNPATFPLNATLYADFSDHNTMI 366
Db 301 PGYDVRSVGNANFNASVKLKESEYQDQ-----KWLSTHDTIDIL 342
QY 367 SIFPALGLYNGTLPSTSTVESIEEDGYASNTVTPFAARAYVEMQCAKEPLRVLV 426
Db 343 NYLTIGIIDQNNMLTAERHVPFMENT--FHRSWVYVQGARVYTEKFCQ--SNDTYVRYVI 398
QY 427 NDRVVLHGCAVDKLGCRKRDDE-----VEGLSTAR 457

Db 399 NDAVPLETCTGTGGSCSEINDEYGYAEKRVAGTDFLK 436
RESULT 6
PPA5_YEAST
ID PPA5_YEAST STANDARD; PRT; 467 AA.
AC P00635;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2) (P60).
GN PHO5 OR YBR093C OR YBR0814.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RX SEQUENCE FROM N.A., AND SEQUENCE OF 18-26.
RX MEDLINE=83168913; PubMed=6300772;
RA Arima K., Oshima T., Kubota I., Nakamura N., Mizunaga T., Toh-E A.;
RT "The nucleotide sequence of the yeast PHO5 gene: a putative precursor
RT of repressible acid phosphatase contains a signal peptide.";
RL Nucleic Acids Res. 11:1657-1672(1983).
RN [2]
RX SEQUENCE FROM N.A., AND SEQUENCE OF 18-45.
RX MEDLINE=85037940; PubMed=6093051;
RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
RT "Structural analysis of the two tandemly repeated acid phosphatase
RT genes in yeast.";
RL Nucleic Acids Res. 12:7721-7739(1984).
RN [3]
RX SEQUENCE FROM N.A.
RX STRAIN=S288C;
RX MEDLINE=95208357; PubMed=7900426;
RA Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
RT "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
RL Yeast 10:1363-1381(1994).
RN [4]
RX SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=87064474; PubMed=3537710;
RA Tait-Kamradt A.G., Turner K.J., Kramer R.A., Elliott Q.D.,
RA Bostian S.J., Thill G.P., Rogers D.T., Bostian K.A.;
RT "Reciprocal regulation of the tandemly duplicated PHO5/PHO3 gene
RT cluster within the acid phosphatase multigene family of Saccharomyces
RL cerevisiae.";
RL Mol. Cell. Biol. 6:1855-1865(1986).
CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC ALCOHOL + ORTHOPHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE.
CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01320; CAA24630.1; -
DR EMBL; X01079; CAA25555.1; -
DR EMBL; X78993; CAA55598.1; -
DR EMBL; Z35962; CAA85046.1; -
DR EMBL; X01080; CAA25556.1; -
DR PIR; S05795; PABYC.
DR PIR; B25241; B25241.
DR HSP; P34752; 1HP.
DR SGD; S0000297; PHO5.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolyase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 103 103
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 456 456
 FT CARBOHYD 461 461
 FT CARBOHYD 461 461
 FT CONFLICT 36 36
 FT CONFLICT 130 130
 FT CONFLICT 294 294
 FT CONFLICT 446 446
 FT CONFLICT 462 462
 FT CONFLICT 466 466
 FT CONFLICT 466 466
 SQ SEQUENCE 467 AA; 52858 MW; DC3C9504BC2D3D0C CRC64;

Query Match 15.9%; Score 392; DB 1; Length 467;
 Best Local Similarity 25.6%; Pred. No. 6.5e-24;
 Matches 123; Conservative 64; Mismatches 216; Indels 78; Gaps 15;

QY 6 VLLSTATFGSTGALPGRNHSCDVGQYCFEISHLWCTSPYFSLADESAISLP 65
 DB 5 VVYSILASLANAGTI--PLGLADVDKIGTKQDIFPL-----GGAPYKSPFGDYGISR 58
 QY 66 DVPDDCRVTFVQVLSRHGARYPTSSAKSAYSALEIAOKNATAFKGYAFLEK-TYNTTLG 124
 DB 59 DLPECEMKQLQWGRHGERYPYTVSLAKTIKSTWYKLSNTRQPNGLSFLNDDYEFIR 118
 QY 125 ADD-----LTPF-GENQVNSGIKFYRYKALKARKIVFP-IPASGSDRVIA 168
 DB 119 DDDLEMETTFANSDVLPNTYTGEMNAKRHRARDFLAQGYGVNENQTSFAVETSNKSRCHD 178
 QY 169 SAEKIEFGQSADKADPGSQHQASFLVNIIV-----PEGSGYNNITLDHGCTAFEDSEL 223
 DB 179 TAQYFIDGL-----GDQ-----FNITLQTVSEAESAGANTLSACNSCPAW-DYDA 222
 QY 224 GDDVEANFTALFAPAIRLEADLPGLVTLTDEDVYVYLMDCPFDTVARTSDATLSPECA 283
 DB 223 NDDIVNEYDTTLDLDAKRLNKENKGLNLTSTDASTLFSWCAFEV-----NAKGYSDVCD 277
 QY 284 LTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFANELLARLTHSPVQDHTSTNHTLD 343
 DB 278 IFTKDELVHYSYQDLHTYHYEGPGYDITKSQVSNLFNASVKLLKQSEIQDQ----- 329
 QY 344 SNPATFPLNATLYADFSHNTMISITFFALGLYNGTKPLSTTSVESIETDGYASWTVPF 403
 DB 330 -----KWLVSFTHTDILNLTAGIIDDKNLTAEVVPMGNT--FHRSWYVPQ 377
 QY 404 ARAYVEMQCAQAEPLRVLVNDVPLHCCAVDKLGRCKRDF-----VEGLSFA 456
 DB 378 GARVYTEKFCQ--SNDTYVRYVINDAVVPIETCTSGPGFSGEINDFYDAEKRVAGTDFL 435
 QY 457 R 457
 DB 436 K 436

RESULT 7
 PPA3_YEAST
 ID PPA3_YEAST
 AC P24031;

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO3 OR YBR092C OR YBR0813.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 OX [1]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=85037940; Pubmed=6093051;
 RX Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RA "Structural analysis of the two tandemly repeated acid phosphatase
 RT genes in yeast.";
 RL Nucleic Acids Res. 12:7721-7739(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=S288C;
 RC MEDLINE=95208357; Pubmed=7900426;
 RX Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
 RA "Analysis of a 70 kb region on the right arm of yeast chromosome II.";
 RL Yeast 10:1363-1381(1994).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X01080; CA25557.1; -
 DR EMBL; X78993; CA25557.1; -
 DR EMBL; Z35961; CA25557.1; -
 DR PIR; S44674; PABYCC.
 DR HSP; P34752; 1IHP.
 DR SGD; S0000296; PHO3.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat_1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolyase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 17
 FT CHAIN 18 467
 FT ACT_SITE 75 75
 FT ACT_SITE 337 337
 FT CARBOHYD 97 97
 FT CARBOHYD 103 103
 FT CARBOHYD 162 162
 FT CARBOHYD 192 192
 FT CARBOHYD 250 250
 FT CARBOHYD 315 315
 FT CARBOHYD 356 356
 FT CARBOHYD 390 390
 FT CARBOHYD 439 439
 FT CARBOHYD 445 445
 FT CARBOHYD 456 456
 FT CARBOHYD 461 461
 FT CONFLICT 219 221
 FT CONFLICT 461 461
 SQ SEQUENCE 467 AA; 52776 MW; 05FEB80DEB41B0FF CRC64;

Query Match 15.4%; Score 380; DB 1; Length 467;
 Best Local Similarity 26.3%; Pred. No. 5.9e-23;
 Matches 114; Conservative 61; Mismatches 196; Indels 62; Gaps 11;

QY 49 GTYSPYFSLADESAISPDVDCRVTFVQVLSRHGARYPTSSAKSAYSALEIAOKNATA 108
 DB 42 GGAGPYSPFGDYGISRDLPEGCEMKQLQWGRHGERYPYTVSKGATIMKTWYKLSNTRQ 101

QY 109 FKGYAFK-TYNYTLGADD-----LTPF-GENOMVNSGKIFRYRRYKALAR 152
 Db 102 FNGSLFLNDDVEFFIRDDDDLEMETTEFANSNVLPYTGEMDAKHAREFLAQGYMEE 161
 QY 153 KIVPF-IRASGSDRVTAKEFIEGFSQAKLADPGSQPHQAFVINVIIPEGSGYNNITLD 211
 Db 162 NOTSFPIFAASSERHDTAQYFIDGL-----GDFNISLQTVSEAMSAGA---NTLS 210
 QY 212 HGCTAFEDSELGDVEANFTALFAPAIRARLEADLPVTLTDEOVVYLMDCPFDTVAR 271
 Db 211 AGNACPGWDEDNDILDKYDITYLDDIAKLNKENKGNLTSKDANTLFAWCAYELNAR 270
 QY 272 TSDATFELSPCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSP 331
 Db 271 -----GYSDVCIDFTEDELVRYSYQDLVSVFYQDGPYDMIRSVGANLFNALKLQKSE 325
 QY 332 VODHSTNHTLDSNPATPLNATLVADFSHDMTMSIFFALGLYNGTKPLSTTSVESIEE 391
 Db 326 TOD-----LKWLSFTHTDILNYLTAGTIDDKNNLTAEYVPFMGN 367
 QY 392 TDGYSASVTPFAARAYVEMMOCQAEKEPLVRVLYNDVVRVPLHGCVAVDKLGCRKDDF-- 449
 Db 368 T--FKSWTVPOGARVYTEKFC--SNDYVVRVINDAVVPIETCTSGPGFCEINDFYD 423
 QY 450 -----VEGLSFAR 457
 Db 424 YAEKRVAGTDFLK 436

RESULT 8
 PPAL_PICPA
 ID PPAL_PICPA STANDARD; PRI; 468 AA.
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE ACID PHOSPHATASE PHO1 PRECURSOR (BC 3.1.3.2).
 GN PHO1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN [1]
 RP MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E.; Gannon P.M.; Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 characterisation of the gene and its product.";
 RL Gene 163:19-26(1995).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 ALCOHOL + ORTHOPHOSPHATE.
 CC -!- INDUCTION: BY PHOSPHATE STARVATION.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; U28658; AAA85503.1; -
 DR HSP; P34752; IHP.
 DR InterPro; IPR000560; -
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 468 ACID PHOSPHATASE PHO1.
 FT ACT_SITE 84 84 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).

FT ACT_SITE 345 345 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 52690 MW; AE5558E27718C2C CRC64;

Query Match 14.4%; Score 355; DB 1; Length 468;
 Best Local Similarity 25.4%; Pred. No. 6e-21;
 Matches 125; Conservative 79; Mismatches 199; Indels 90; Gaps 21;

QY 6 VLLSTATL---FGSTSGTALPGRNSHSCDITVDGGYQCPEISHLWGTYSFY--PSLADE 60
 Db 10 IILALATQSVFAVELQHVLVGNDRPYPORT-DDQYNILRLHGLGL-GPYIGYNGWIAAE 67
 QY 61 SAISPDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIE-AIQKNATAFKGYAFKTV 119
 Db 68 SEI-----ESCIDQAHLLMRHGERYPSTNVGKQLEALYQKLLDADVEVPTGPLSEFQDY 122
 QY 120 NYTLG-----ADLTP---FGENQMVNSGKIFRYRYKAL-----ARKIVPFFIRASGSDR 165
 Db 123 DYFVSDAANYEQETTKGYFSGLTAFDFTLRLKRYVDHLINTSECKKL--SVWAGSQER 180
 QY 166 VIASAEKFTIEGFSQAKLADPGSQPHQAFVINVIIPE---GSGYNNITLDHGTC---A 217
 Db 181 VYDTAKYFAQGMKSNYTD-----MVEVVALEEEKSQGLNSLTARISCPNYNSHI 230
 QY 218 FEDSELGDDV---EANFTALFAPAIRARLEADLPVTLTDEOVVYLMDCPFDTVARTSD 274
 Db 231 YKDGFPNDIAERAD-----RLNTLSPGENITADDIPTIYALYCGGELNVRGE- 278
 QY 275 ATELSPFCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVOD 334
 Db 279 ----SSFCDLVSREALYAYLRDLGWYVNGVNGPLGTIGYVYAN----- 321
 QY 335 HTSTNHTLDSNPATPLNATLVADFSHDMTMSIFFALGLYNGTK-PLSTTSVESIEETD 393
 Db 322 --ATROLLENTAD-PRDYPLYFSFSDHLDLQVFTSLGLFNVDLP-----DQIQFQT 373
 QY 394 GYSASVTPFAARAYVEMMOC--QAEKEPLVRVLYNDVVRVPLHGCVAVDKLGCRKDDVE 451
 Db 374 SFKSTFIVPMGARLLTERLLCHVEGEKYVVRILNDVAVPLSDCSSGGFGSCPLNDIYS 433
 QY 452 GLSFARSGGNWAE 464
 Db 434 RLEALNEDSDFAE 446

RESULT 9
 PPAD_YEAST
 ID PPAD_YEAST STANDARD; PRI; 468 AA.
 AC P52290;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE PROBABLE ACID PHOSPHATASE YDL024C PRECURSOR (EC 3.1.3.2).
 GN YDL024C OR D2815.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RA Andre B.; Visser S.; Urrestarazu L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DDBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z48432; CA88335.1; -;
 DR EMBL; Z74072; CA98583.1; -;
 DR SGD; S0002182; YDL024C.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 468
 FT ACT_SITE 76 76 PROBABLE ACID PHOSPHATASE YDL024C.
 FT ACT_SITE 338 338 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT CARBOHYD 98 98 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 462 462 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 53076 MW; 1F07FF374DDF162C CRC64;

Query Match 14.4%; Score 354; DB 1; Length 468;

Best Local Similarity 25.9%; Pred. No. 7.2e-21;
 Matches 114; Conservative 63; Mismatches 183; Indels 80; Gaps 14;

QY 49 GTYSPYFLADESAISPOVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATA 108
 Db 43 GGSAPYFPPANYGIPTDIPEGCRETQVOMIGRHGERTPTREAKDIFVWYKISNTGK 102
 QY 109 FKGKFAFLKT-YNYTL-----GADLTTP-GENQVNSGKRYRYKALAR 152
 Db 103 YEGSLFSLNNGYEFFIPDESLEMETTLQNSIDVLPYTGEMNAKRHAREFLAKYGLME 162
 QY 153 KIVPF-IRASGSDRVIAAEKFIE---GFOSA--KLADPGSQPHQASPVINVIPEGS 205
 Db 163 NCTNPFITTSKRIYDTAQVFAEALGDFNLSLOTISENS-----SG 206
 QY 206 YNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCP 265
 Db 207 ANTLAAKSCPNW-NSMANNDILMSYSDYLENISDRLDENKGLNSRKDAALFSCWA 265
 QY 266 FTVARTSDATSELSFCALFTHDEWIQDYLOSGLKYYGAGNPLGPAQGVGFAN---E 322
 Db 266 FEL-----NAKGYSNICDIFSAAELIHTSYETDLTSTFYQNGPGYKLIKISGANLFNATVK 320
 QY 323 LIARLTHSPVQDHTSTNTHLDSNPATFPFLNATLYADFSDHNTMISIFPALGLNGTKPLS 382
 Db 321 LIHQSAH-----LDQKVLSTFTHDDIILNYLTAGLIDDRNL 359
 QY 383 TTSVESIEETDGYASWTVPFAARAYVEMMQCAEKEPLVRLVNDRVVPLHGCADVCKLG 442
 Db 360 TNHVFPDRHS--YHRSWYIPOGARVYTEKFQC--SNDSYRVYVYVNDVAVPIESCSPGPF 415
 QY 443 RCKRDFVE-----GLSF 455
 Db 416 SCEEGTFVEYKDRLRGVSF 435

RESULT 10
 PHYB_ASPNG

ID PHYB_ASPNG STANDARD; PRT; 479 AA.
 AC P34754;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHYB.
 GN Aspergillus niger.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.
 RX MEDLINE=93371452; PubMed=7916610;
 RA Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,
 RA Ullah A.H.J.;
 RT Identification and cloning of a second phytase gene (phyB) from
 RT Aspergillus niger (ficusum).";
 RL Biochem. Biophys. Res. Commun. 195:53-57(1993).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTASE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).

Query Match 14.2%; Score 348.5; DB 1; Length 479;
 Best Local Similarity 29.3%; Pred. No. 2e-20;
 Matches 119; Conservative 55; Mismatches 169; Indels 63; Gaps 17;

QY 49 GTYSPYFLADESAISPOVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNAT 107
 Db 53 GPKYSERVSY----GIARDPTGCEVDQVIMVKRHGERYPSPSAGKSIEALAKVYSINT 108
 QY 108 AFKGKFAFLKTYNYTL-----GADLT-PP-GENQVNSGKIFRYRYKAL---ARKIYVP 157
 Db 109 EYKGDIAFLNDWTYYV 168
 QY 158 IRASGSDRVIAAEKFIEGFSQAKLADPGSQPHQASPVINVII-PEGSGYNTLDHGTCT 216
 Db 169 F-SSGGRVETARKFEGGFFGN-----YSTNAALNIISEVVMGADSLTP--TCD 217

QY 217 AFDSSELGDDVEANFTALFAPAIRARLEADLPGLVTLTDEDVYVLMDCPFDTVARTSDAT 276
 Db 218 TDNDQTTCDNLTYQLPQFVAA--ARLNSQNGMNLTSADVTLNLMVMAFELNAR----- 270
 QY 277 ELSPFCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGGVGFANELIARLTHSPVDHT 336
 Db 271 PFSNNINAFQTDEWYSGFYVEDLNYCYCAGPGDKKMAAVGAVYANASLTILNQGP----- 325
 QY 337 SINHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLY--NGTKPLSTTSVESIEETDG 394
 Db 326 -----KEAGP-----LFFNFADHTNITPILAAAGLVLPNEDLPL-----DRVAFGNP 367
 QY 395 YSASWTVPFAARAYVEMMOCA-----EKEPLVRLVNDVRVPLHGC 436
 Db 368 YSIGNIVPMGGHLTIERLSQATLSDAGTYVRLVINEAVLPFND 413

Query Match 14.1%; Score 347.5; DB 1; Length 479;
 Best Local Similarity 29.1%; Pred. No. 2.5e-20;
 Matches 118; Conservative 54; Mismatches 171; Indels 63; Gaps 16;

RESULT 11
 PHYB_ASPAW
 ID PHYB_ASPAW STANDARD; PRT; 479 AA.
 AC P34755;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 33, Last annotation update)
 DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
 DE 3-PHOSPHOHYDROLASE B) (PH 2.5 OPTIMUM ACID PHOSPHATASE).
 GN PHYB OR APH.
 OS Aspergillus awamori.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=105351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ALK0243;
 RX MEDLINE=94040796; PubMed=8224894;
 RA Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
 RA Miettinen-Oinonen A., Nevalainen H., Rambosk J.A.;
 RT "The cloning and sequencing of the genes encoding phytase (phy) and
 RT pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
 RT awamori".
 RL Gene 133:55-62(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=99264417; PubMed=10329192;
 RA Kostrewa D., Wyss M., D'Arcy A., van Loon A.P.;
 RT "Crystal structure of Aspergillus niger pH 2.5 acid phosphatase at
 RT 2.4-A resolution".
 RL J. Mol. Biol. 288:965-974(1999).
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE.
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
 CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 CC EMBL; LO2420; AAA16897.1;
 CC PIR; JN0890; JN0890.
 CC PDB; 1QFX; 19-APR-00.
 CC InterPro; IPR000560;
 CC Pfam; PF00328; acid_phosphat_1.
 CC PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 CC PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 CC Hydrolase; Glycoprotein; Signal; 3d-structure.
 KW
 FT SIGNAL 1 19 BY SIMILARITY.

FT CHAIN 20 479 3-PHYTASE B.
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR.
 FT ACT_SITE 337 337 PROTON DONOR.
 FT DISULFID 71 387
 FT DISULFID 128 472
 FT DISULFID 216 441
 FT DISULFID 225 298
 FT DISULFID 413 421
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 479 AA; 52678 MW; 4F8E0F3778CCB08 CRC64;

Query Match 14.1%; Score 347.5; DB 1; Length 479;
 Best Local Similarity 29.1%; Pred. No. 2.5e-20;
 Matches 118; Conservative 54; Mismatches 171; Indels 63; Gaps 16;

QY 49 GTYSPYFSLADESALSIPDVPDDCRVTFVQVLSRHGARYPTSSASK--AYSALIEATQKNAT 107
 Db 53 GYSERVSY-----GIARDPPTSCVDQVIMVKRIGERYPSAGKDIEELAKVYSINTT 108
 QY 108 AFGKVAFLKTYNYTL-----CADDLT-PF-CENQWVNSGKIFRYRYKAL--ARKIVPF 157
 Db 109 EYKGDIAFLNDWTYYVNECYNAETSGPYAGLLDAYNHGNDYKARYGHLNGETVVPF 168
 QY 158 IRASGSDRVIASAEKFIIEGFQSAKLADPGSQPHQASPVINVII-PEGSGYNNNTLDHGTC 216
 Db 169 F-SSGVRVETARKFEGEGFGYN-----YSTNAALNIISEVVGADSLTP--TCD 217
 QY 217 AFDSSELGDDVEANFTALFAPAIRARLEADLPGLVTLTDEDVYVLMDCPFDTVARTSDAT 276
 Db 218 TDNDQTTCDNLTYQLPQFVAA--ARLNSQNGMNLTSADVTLNLMVMAFELNAR----- 270
 QY 277 ELSPFCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGGVGFANELIARLTHSPVDHT 336
 Db 271 PFSNNINAFQTDEWYSGFYVEDLNYCYCAGPGDKKMAAVGAVYANASLTILNQGP----- 327
 QY 337 SINHTLDSNPATFPLNATLYADFSHDNTWISIFFALGLY--NGTKPLSTTSVESIEETDG 394
 Db 326 -----AGSLFFNFADHTNITPILAAAGLVLPNEDLPL-----DRVAFGNP 367
 QY 395 YSASWTVPFAARAYVEMMOCA-----EKEPLVRLVNDVRVPLHGC 436
 Db 368 YSIGNIVPMGGHLTIERLSQATLSDAGTYVRLVINEAVLPFND 413

RESULT 12
 PPA5_KLULA
 ID PPA5_KLULA STANDARD; PRT; 469 AA.
 AC P52289;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHOS.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RA Fermanian E.;
 RL Thesis (1995), University of Salamanca, Spain.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; 233995; CAA83964.1; --
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 16
 FT CHAIN 17 469
 FT ACT_SITE 77 77
 FT ACT_SITE 339 339
 FT CARBOHYD 23 23
 FT CARBOHYD 31 31
 FT CARBOHYD 32 32
 FT CARBOHYD 129 129
 FT CARBOHYD 201 201
 FT CARBOHYD 229 229
 FT CARBOHYD 250 250
 FT CARBOHYD 317 317
 FT CARBOHYD 392 392
 FT CARBOHYD 447 447
 FT SEQUENCE 469 AA; 52532 MW; 863B528D0740AA7E CRC64;

Query Match 13.9%; Score 341.5; DB 1; Length 469;
 Best Local Similarity 24.5%; Pred. No. 7.2e-20;
 Matches 121; Conservative 75; Mismatches 202; Indels 95; Gaps 17;

QY 7 LLSIATFGSTGALPGNCHSDVDGYYQCF-----PRLSHLWGTSPYFSLA 58
 DB 1 MSLILGLLSLSTHAAPIK-----DNGTVCYALNNSTTDESIFSLNGQCPHYDYP 53
 QY 59 DESAIPDVPDCRVTFVQVLSRHGARYPTSSAKAYASALTEAIOKNATAPKGYAFLEKT 118
 DB 54 QSGFIPVEPVDCTVHVQVQLARHGERIPTASKGLMIALWDKLE-----FGQY----- 104
 QY 119 YNYTLGADDLTPFGNQVNSGIKYYR-----KALARKIVPFIR----- 159
 DB 105 -----NDPLEVNDYEFVSNKYFDQLTNDVDPSNPYAGAKTAQHLGKYIAYNGD 158
 QY 160 -ASGSDRVTAESAETEGTQSAKLADPGSQPHQASPVINVIPEG--SGYNNLTDHCTCT 216
 DB 159 LFSDSNPVFTSSGRVH--QTAKYVYSSLEELDQLDQIIQENETSGANSLTADSCM 216
 QY 217 AFEDSELGDDVEANFTALFAPAIRAR---LEADLPQVTLTDEDVYLMDCMPCFTVARTS 273
 DB 217 TY-NGDLGEYFENATLPYLTIDKNWKKNSNL-NITLHDDIELLDVCAFEINRGS 274
 QY 274 DATLSPPCALFTHDEWIDYDQLSQIGKYGAGNPLGPAQGVGFANELIARLTHSPVQ 333
 DB 275 SAV-----CDLPERNDLVAYSANNYNNRRGAGNPMNPGSVLVNAYSNLLTQADEL 329
 QY 334 DHTSTNHTLDSNPATPLNATLYADFSHONTMISIFFALGLY-NGTKPLSTTSVESIET 392
 DB 330 DN-----KWLSESHDDTDIQOFISALGLDING-----VTEISLQGVDFQ 368
 QY 393 DGYSASVTPVFAARAYVEMQCOAEKPELVRLVNDVRVYPLHGCAVDKLGCRKDDF--- 449
 DB 369 NIQQLSNVTPMGGRIFTEKLC--GNASVRYIINDVILVVPVGTSGPGFSCPIEDFDY 426
 QY 450 -----VEGLSFARS 458
 DB 427 ITNRLNGIDYVSS 439

RESULT 13
 PPAL_SCHPO

ID PPAL_SCHPO STANDARD; PRT: 453 AA.
 AC P08091.
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-2000 (Rel. 40, Last annotation update)
 DE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO1 OR SPBP4G3.02.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86140050; PubMed=3005272;
 RA Elliott S., Chang C., Schweingruber M.E., Schaller J., Rickli E.E.,
 RA Carbon J.;
 RT "Isolation and characterization of the structural gene for secreted
 RT acid phosphatase from Schizosaccharomyces pombe.";
 RL J. Biol. Chem. 261:2936-2941(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Rieger M., McDougall R.C., Rajandream M.A., Barrell B.G.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: CELL WALL.
 CC -1- INDUCTION: REPRESSED BY PHOSPHATE AND WEAKLY BY THIAMINE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; M11857; AAA35321.1; --
 EMBL; AL137099; CAB68657.1; --
 PIR; A25326; A25326.
 HSP; P34752; JIHP.
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 453
 FT ACT_SITE 69 69
 FT ACT_SITE 330 330
 FT CARBOHYD 95 95
 FT CARBOHYD 151 151
 FT CARBOHYD 183 183
 FT CARBOHYD 193 193
 FT CARBOHYD 243 243
 FT CARBOHYD 319 319
 FT CARBOHYD 410 410
 FT CARBOHYD 429 429
 FT CARBOHYD 443 443
 FT SEQUENCE 453 AA; 50557 MW; 7CF891256BE154D1 CRC64;

Query Match 12.1%; Score 297; DB 1; Length 453;
 Best Local Similarity 24.8%; Pred. No. 2.5e-16;
 Matches 99; Conservative 60; Mismatches 190; Indels 50; Gaps 9;

QY 51 YPSFSLADESAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYAS-----LIE 100
 DB 43 HKPYF-----YGPSIDPPTCKIKQVHTLQHGSRNPT--GNAADFADGIANFQORLLG 96
 QY 101 AIQKNATAFKGYAFLEKTYNTL---GADDLTPFGNQVNSGIKYYRKYKALARKIVPF 157

Db 97 SVPIDYVSNGNLSFVPTWPIVEAANADALSSGRVLFDMGROFYERYHELNFNASTYN 156
 QY 158 IRASGSDRVIAAEKFIQFQSAKLADPGSQPHQASPVINVIIPGSGYNNLTDHGTCTA 217
 Db 157 IYTAQQRVDSALWYGMF-----GEDVHNFTNTILVSENATAGSNLSYSSNACPA 209
 QY 218 FEDSELGDDVEANFTALFAPALRLEADLPQVTLTDDVVYLMDCPFDVARTSDATE 277
 Db 210 SDADFTTALAEARNVMPYRQRLNPFYSNNTNDIINLYGICSEYETALQ-----D 264
 QY 278 LSPFCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTS 337
 Db 265 YSEFCLENSVDLFNEVEGDLFSYGMGNSVKWGSIFGGAYANSLNSL--RSVENNTQ 322
 QY 338 TNHTLDSNPATPLNATLYADFSHNTMISIFFALGLNGTKPLSTTSVESIETDGVSA 397
 Db 323 -----QVFAFTHDANIIPVETALGFFTDNTPENPLPSTSYQVHSHSKA 366
 QY 398 SWTVFAARAYVEMMQCOAEKEPLVRLVNDKRVPLHGC 436
 Db 367 SEFVPFAGNLITELQCEDSKY-YVRLVNEEVPLSDC 404

RESULT 14
 ID PPA2 SCHPO STANDARD; PRT; 463 AA.
 AC Q01682;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE THIAMINE-REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PH04 OR SPBC428.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-30.
 RX MEDLINE=91064763; PubMed=2249257;
 RA Yang J., Schweingruber M.E.;
 RT "The structural gene coding for thiamin-repressible acid phosphatase
 in Schizosaccharomycetes pombe.";
 RL Curr. Genet. 18:269-272(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Rajandream M.A., Barrell B.G., Baker S., Mungall K.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY DEPHOSPHORYLATE THIAMINE PHOSPHATES.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: CELL WALL.
 CC -!- INDUCTION: RERESSED BY THIAMINE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

 DR EMBL; X56939; CAAG0258.1; --
 DR EMBL; AL034382; CAAG22278.1; --
 DR PIR; S14119; S14119.
 DR InterPro; IPR000560; --
 DR Pfam; PF00328; acid.phosphat.1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Cell wall; signal.

FT SIGNAL 1 18
 FT CHAIN 19 463
 FT ACT_SITE 69 69
 FT PROTON_DONOR (BY SIMILARITY).
 FT CARBOHYD 98 98
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 104 104
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 221 221
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 251 251
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 328 328
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 433 433
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 439 439
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 458 458
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 463 AA; 52118 MW; F48DAFF8BB6B234A CRC64;

Query Match 11.3%; Score 278.5; DB 1; Length 463;
 Best Local Similarity 25.5%; Pred. No. 7.8e-15;
 Matches 108; Conservative 65; Mismatches 187; Indels 63; Gaps 15;

QY 46 HLWGTVS-----PYFSLADESAISPDPVDDCRVTVFQVLSRHGARYPT--SSASKAYS-- 97
 Db 35 HL-GTISVYHEPYF-----NGPTTSPESCAIKQVHLLQRHGSRNPTGDDTATDVSSAQY 88
 QY 98 -----LLEATQKNATAFKGYAFKTYNTL---GADLTFFGNGQMVNSGKIFYRR 146
 Db 89 IDIFONKLLNGSIPVNFSPENPLYFVKHWTPVKAENADQLSSSGRIELFDLGRQVFER 148
 QY 147 YKALARKIVPFIASGSDRVIAAEKFIQFQSAKLADPGSQPHQASPVINVIIP--GS 204
 Db 149 YLELFDVDVYDIINTAAQVRVDSAEWFSYCMFGDDMQNKTN-----FVLPEDDSA 199
 QY 205 GYNTLIDHGTCTAFEDSELGDD--VEANFT---LFPATRARLEADL-PGVILTDBDVVY 259
 Db 200 GANSLAMYSCPVYEDNNIDENTEAAHTSRNVFLKPIANRLNKYFDGYNLTVDVRS 259
 QY 260 LMDMCPFDVTARTSDATELSPFCALFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGF 319
 Db 260 LYICYVETALRDN-----SDFCSLTFPSEFLNFEVDSDLDVAYWGGPASEWASTLGGAY 314
 QY 320 ANELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFSHNTMISIFFALGLYNGTK 379
 Db 315 VNNLANL-----RKGVNNASDRK-----VELAFTHDSQIIPVEAALGFPPDIT 358
 QY 380 PLSTTSVESIETDGVSAKSWTFFAARAVEMMQCOAEKEPLVRLVNDKRVPLHGCAYD 439
 Db 359 PEHLPTDKNIFTYSLKTSFSFVFAAGNLITELFLC--SDNKYYVRLVNDKRVPLHGCAYD 417
 QY 440 KLG 442
 Db 418 PSG 420

RESULT 15
 ID PPA2 CAEL STANDARD; PRT; 755 AA.
 AC Q09549; Q17843;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE ACID PHOSPHATASE F26C11.1 (EC 3.1.3.2).
 GN F26C11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Matthews P., Lloyd C.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 247072; CAA87370.1; -;
DR EMBL; 254342; CAA87370.1; JOINED.
DR EMBL; 254342; CAA91156.1; -;
DR EMBL; 247072; CAA91156.1; JOINED.
DR WormPep; F26C11.1; CE05732.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein; Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
SQ SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.38; Score 130; DB 1; Length 755;
Best Local Similarity 21.9%; Pred. No. 0.012;
Matches 90; Conservative 55; Mismatches 130; Indels 136; Gaps 22;

QY 57 LADESAISPDVDDCRVTFVOVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFL 116
DB 332 LAIEKSFSEV-DNMKLEFVOTIWRHGD-----SALEGLFPISEKNWTFGG----- 378
QY 117 KTYNTLGAADLTTPGEMQVNSGIKYRYY-----KALARKIVPFIRASGSDRVI 167
DB 379 -----GLGELTPMGSEMNNLGTIFRRYVEDQQFLSHRYAAKET--YIRSTNLNRTI 429
QY 168 ASAEKFIQFQSAKLADPGSQPHQASPVINVIIEGSGYNNLTLD--HG----- 213
DB 430 ISAMSLLYG-----MFPPGAW-----NI---OGVDYPNDVDVMDQGGTFIPVHVDGID 473
QY 214 -----TCTAFED-----SELGDDVEANFTALFAIRARLEADLPVTLTDEDVYVL 260
DB 474 QCAVAQLCNCRRFOLEKWAEL-DEVKNATVAMI--ALNRRVAA---FYNVTDQP----- 523
QY 261 MDMCPFDVARTSDATELSPPC-----ALFTHDEWIDYDYLQSLGKYG-----YG 306
DB 524 -----EKENRYTDAWK-----CORNWFNDTMYQQLPWYNEDLYNEAORTYAPKRRFTEGN 573
QY 307 AGNP-----LGPAGGVGFANELIARLTHSPVQDHTSTNHTLD--SNPATFPLN 352
DB 574 FGNRPSPVGDIDIPQEVSTLQGGPLNLEIFEREKIRCVADAEKNSIDVLPKLF--- 630
QY 353 ATLYADFSDHNTMISIFPGLGYNKTRPLSTSVESIEETDGYGASWTVPF 403
DB 631 ---YAYSSHDQLVYALLVTLGI-----TDVVKTVDGWPDTSSSLTIEY 670

Search completed: October 26, 2001, 15:18:04
Job time: 450 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:04 ; Search time 57.24 Seconds
(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGFFVLLSIATLFGSTGCT.....DEVEGLSFARSGNWAECFA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2462	100.0	467	20 AAY43169	Consensus phytase-
2	2462	100.0	467	21 AAB20526	Consensus phytase
3	2462	100.0	467	21 AAY69568	Mutant phytase-1,
4	2443	99.2	467	21 AAB20532	Consensus phytase
5	2438	99.0	467	21 AAB20531	Consensus phytase
6	2414	98.1	467	20 AAW93382	Fungal consensus p
7	2408	97.8	467	20 AAW93380	Fungal phytase pro
8	2408	97.8	467	20 AAW93381	Fungal consensus p
9	2408	97.8	467	21 AAB20515	Consensus phytase
10	2408	97.8	467	21 AAY69558	Phytase-1, a conse
11	2407	97.8	467	20 AAW93383	Fungal consensus p

12	2405	97.7	467	20 AAW93384	Fungal consensus p
13	2399	97.4	467	20 AAW93385	Fungal consensus p
14	2398	97.4	467	20 AAY39906	Ascomycete Consens
15	2362	95.9	467	20 AAY43170	Consensus phytase-
16	2362	95.9	467	21 AAB20527	Consensus phytase
17	2362	95.9	467	21 AAY69569	Mutant phytase-10,
18	2360	95.9	467	21 AAB20534	Consensus phytase
19	2355	95.7	467	21 AAB20533	Consensus phytase
20	2342	95.1	467	21 AAB20524	Consensus phytase
21	2342	95.1	467	21 AAY69566	Phytase-10, a cons
22	2290	93.0	467	20 AAY43172	Consensus phytase-
23	2290	93.0	467	21 AAY69571	Phytase-7, a deriv
24	2287	92.9	467	21 AAB20529	Consensus phytase
25	2282	92.7	441	21 AAB20514	Consensus phytase
26	2216	90.0	441	21 AAB20523	Consensus phytase
27	2215	90.0	467	21 AAB20530	Consensus phytase
28	2215	90.0	467	21 AAY69572	Consensus phytase
29	2134	86.7	431	21 AAB20535	Consphyl2, a deriv
30	2134	86.7	431	21 AAY69557	Consensus phytase
31	2044	83.0	437	21 AAB20525	Initial consensus
32	2044	83.0	437	21 AAY69567	Consensus phytase
33	2032	82.5	467	21 AAY69574	Phytase-11, a cons
34	2012	81.7	467	20 AAY43171	Mutant Aspergillus
35	2012	81.7	467	21 AAB20528	A. fumigatus phyta
36	2012	81.7	467	21 AAY69570	Aspergillus fumiga
37	1956.5	79.5	424	21 AAB20536	Mutant Aspergillus
38	1956.5	79.5	424	21 AAY69565	Consensus phytase
39	1879	76.3	467	15 AAR46792	Initial consensus
40	1879	76.3	467	15 AAR46234	Phytase. Aspergil
41	1875	76.2	467	12 AAR11333	A. niger phytase.
42	1875	76.2	467	18 AAW15124	Recombinant chromo
43	1875	76.2	467	20 AAY39904	Aspergillus ficum
44	1867	75.8	465	19 AAW84356	A. ficum phytase
45	1867	75.8	465	20 AAY39905	Aspergillus fumiga
					A. fumigatus phyta

ALIGNMENTS

RESULT 1
AAY43169
ID AAY43169 standard; Protein; 467 AA.
XX
AC AAY43169;
XX
XX
DT 06-JAN-2000 (first entry)
XX
DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.
XX
KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
KW consensus sequence.
XX
OS Synthetic.
XX
XX W09948380-AL.
XX
XX
PD 30-SEP-1999.
XX
XX 22-MAR-1999; 99WO-DK00154.
XX
XX 23-MAR-1998; 98DK-0000407.
PR 19-JUN-1998; 98DK-0000806.
PR 18-SEP-1998; 98DK-0001176.
PR 22-JAN-1999; 99DK-0000091.
XX 22-JAN-1999; 99DK-0000093.
XX
XX (NOVO) NOVO-NORDISK AS.
XX
PI Petersen S;
XX
XX WPI; 1999-591030/50.
DR N-PSDB; AAZ31520.
XX

PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 9; 71pp; English.

XX This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 XX The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently. In addition to improved
 CC phytase-expressing transgenic plants, these plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 20; Length 467;
 Best Local Similarity 100.0%; Pred. No. 4e-239;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTYSYFSLADE 60
 DB 1 mgffvllsiatlfgstgaltprgnshscdvtvdggycfpeishlwgtysyfslade 60

QY 61 SAISDPVDDCRVTFVQLSRHGARYPTSSASKAYSALIEALQKNAFAKGYAFLLKTYN 120
 DB 61 saispdvddcrvtfvqlsrhgaryptssaskaysaliealqknafakgyafllkty 120

QY 121 YTLGADDLTPFGENQMVNSGIKFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalar k ivpfir asgsdrv iasae kfi egf qsa 180

QY 181 KLADPGSQHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqhqaspvinviipgsgynnltldhgtctafedselgddveanftalfap 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDVTARTSDATELSPFCALFTHDEWIQYDYLQSLG 300
 DB 241 arleadlpvltltdedvvyv lmdcpfdvtartsd atel spfcal fthdew i qdy lqslg 300

QY 301 KYYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kyygygagnplgpagvgfaneliarlthspvq dhtstnhtl dsnpatf plnatly adfs 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMCOAKEP 420
 DB 361 hdntmisiffalglyngtkplsttsvesieetd gysaw tvpfar ayvemm coakep 420

QY 421 LVRVLVNDRVVPLHGCADVCKLGRCKRDDFVEGLSFARSGGNWAECEFA 467
 DB 421 lrvl vndrvvplhgcadvcklgrckrddfve glsfar sggnwae cfa 467

RESULT 2

AAB20526

ID AAB20526 standard; Protein; 467 AA.

XX AAB20526;

AC AAB20526;

DT 05-DEC-2000 (first entry)

XX Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

XX W0200043503-A1.

PN

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI: 2000-491161/43.

XX N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 21; Length 467;

Best Local Similarity 100.0%; Pred. No. 4e-239;

Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTYSYFSLADE 60
 DB 1 mgffvllsiatlfgstgaltprgnshscdvtvdggycfpeishlwgtysyfslade 60

QY 61 SAISDPVDDCRVTFVQLSRHGARYPTSSASKAYSALIEALQKNAFAKGYAFLLKTYN 120
 DB 61 saispdvddcrvtfvqlsrhgaryptssaskaysaliealqknafakgyafllkty 120

QY 121 YTLGADDLTPFGENQMVNSGIKFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgenqmvnsgikfryrykalar k ivpfir asgsdrv iasae kfi egf qsa 180

QY 181 KLADPGSQHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqhqaspvinviipgsgynnltldhgtctafedselgddveanftalfap 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDVTARTSDATELSPFCALFTHDEWIQYDYLQSLG 300
 DB 241 arleadlpvltltdedvvyv lmdcpfdvtartsd atel spfcal fthdew i qdy lqslg 300

QY 301 KYYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kyygygagnplgpagvgfaneliarlthspvq dhtstnhtl dsnpatf plnatly adfs 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMCOAKEP 420
 DB 361 hdntmisiffalglyngtkplsttsvesieetd gysaw tvpfar ayvemm coakep 420

QY 421 LVRVLVNDRVVPLHGCADVCKLGRCKRDDFVEGLSFARSGGNWAECEFA 467
 DB 421 lrvl vndrvvplhgcadvcklgrckrddfve glsfar sggnwae cfa 467

RESULT 3
 AAY69568
 ID AAY69568 standard; protein; 467 AA.
 XX
 AC AAY69568;
 XX
 DT 19-APR-2000 (first entry)
 XX
 DE Mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.
 XX
 KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutein.
 XX
 OS Aspergillus terreus 9A1.
 OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger T213.
 OS Aspergillus fumigatus str. NRRL3135.
 OS Aspergillus fumigatus ATCC13073.
 OS Aspergillus fumigatus ATCC32722.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Emeritella nidulans.
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Synthetic.
 XX
 EH Key Location/Qualifiers
 FT Peptide 1..26
 FT /note= "Phytase signal peptide from Aspergillus terreus
 FT cbs16.46"
 FT Protein 27..467
 FT /note= "Mature phytase-1-thermo[8]-Q50T-K91A"
 XX
 FN EP969089-AL.
 XX
 PD 05-JAN-2000.
 XX
 XX 23-JUN-1999; 99EP-0111949.
 XX
 PR 29-JUN-1998; 98EP-0111960.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Brugger R, Lehmann M, Wyss M;
 XX
 DR WPI: 2000-099429/09.
 DR N-PSDB; AAZ59715.
 XX
 PT New stabilized enzyme formulation, useful for feed compositions for
 PT monogastric animals -
 XX
 PS Example 5; Fig 19; 101pp; English.
 XX
 CC The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is

CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-1
 CC consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a
 CC temperature optimum and melting point 7 degrees Celsius higher than that
 CC of phytase-1 (AAY69538).
 XX
 XX Sequence 467 AA;

Query Match 100.0%; Score 2462; DB 21; Length 467;
 Best Local Similarity 100.0%; Pred. No. 4e-239;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSLATLFGSTGTALGPRGNHSCDTVDGYSOCFPEISHLWCTYSPYSLADE 60
 DB 1 mgvfvlslatlfgstgtalgrgnshscdtvdgysocfpeishlwctyspyfslade 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdvddcrvtfvqlsrhgaryptssaskaysalieaiqknatafkgyafilktyn 120
 QY 121 YTLGADDLTFPGENQMVNSGKIFRYRYKALARKIVPFIKASGSDRVASAEKFIKGFQSA 180
 DB 121 ytlgaddltfpgenqmvnsgikfryrykalkarkivpfiragsdsviasaekfiegfqsa 180
 QY 181 KLADPGSQPHQASPVINVIPEGSGYNNTLDHGTCTAPEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsqphqaspvinvipegsgynntldhgtctaredselgddveanftalfapair 240
 QY 241 ARLEADLPVITDEVDVYLMDCPFDTVARTSDATELSPFCALFTHDEWIQDYLSIG 300
 DB 241 arleadlpvltldevdvylmcmcpfdtvarstatedelspfcalthdewiqdylsig 300
 QY 301 KYGYGAGNPLGPAOCVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpacgvgfaneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMQCAKEP 420
 DB 361 hdntmisiffalglyngtkplsttsvesieetdgyasawtvpfarayvemmqcakep 420
 QY 421 LVRVLVNDRVPLHGCANDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndrvplhgcavdklgrckrdddfveglsfarsggnwaecefa 467

RESULT 4

AAB20532
 ID AAB20532 standard; Protein; 467 AA.

AC AAB20532;

DT 05-DEC-2000 (first entry)

XX Consensus phytase 3 thermo 11 Q50T K91A protein SEQ ID NO:93.

DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

XX WO2000/43503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.
 PR 21-SEP-1999; 99DK-0001340.
 XX
 XX (NOVO) NOVO NORDISK AS.
 XX
 PI Lehmann M;

XX WPI; 2000-491161/43.
 DR N-PSDB; AAA73291.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 23a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 99.2%; Score 2443; DB 21; Length 467;
 Best Local Similarity 99.4%; Pred. No. 3.3e-237;
 Matches 464; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGPRGNSHSCDVTGGYOCFPEISHLWCTYSPYFSLADE 60
 DB 1 mgvfvlslatlfgstgaltgprgnshscdvtggycfpeishlwctyspyfslade 60
 QY 61 SAISPDPVDCRVTFFQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdpvdkcrvtfvqlsrhgaryptssaskaysalialieaiqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgnqmvnsgikfrryrykalkarkivpfirasgsdrviasaekfiqfqs 180
 QY 181 KLADPGSOPHOASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsophoaspvinviipgsgynntldhgtctafedstlgddveanftalfapair 240
 QY 241 ARLEADLPGLTDEDDVYVLMDCPFDVTARTSDATLSLSPFCALFTHDEMIOYDYLQSLG 300
 DB 241 arleadlpgltdeddvylmldcpfdvtartsdatslspfcalfthdewiqdyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVDHTSTNHTLDSNPATPLNATLYADES 360
 DB 301 kygygagnplgpaqgvgfaneliarlthspvqdhtstnhtldsnpatplnatlyads 360
 QY 361 HDNTMISIFFALGLYNTKPLSTTSVESIETDGYASWTVPFAARAYVEMMOCQAEKEP 420
 DB 361 hdnTmisiffalglyntkplsttsvesietdgyaswtvpfaarayvemmocqaekep 420
 QY 421 LVRVLVNDVRVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 lvrvlvndvrvpvhgcavdkgcrkddfveglsfarsggnwaecefa 467

RESULT
 AAB20531

ID AAB20531 standard; Protein; 467 AA.

XX AC AAB20531;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX OS Synthetic.

XX PN WO2000043503-A1.

XX PD 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-DK00025.

XX PR 22-JAN-1999; 99DK-0000092.

XX PR 21-SEP-1999; 99DK-0001340.

XX PA (NOVO) NOVO NORDISK AS.

XX PI Lehmann M;

XX DR WPI; 2000-491161/43.

XX DR N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Disclosure; Fig 22a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX SQ Sequence 467 AA;

Query Match 99.0%; Score 2438; DB 21; Length 467;
 Best Local Similarity 99.1%; Pred. No. 1e-236;
 Matches 463; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALGPRGNSHSCDVTGGYOCFPEISHLWCTYSPYFSLADE 60
 DB 1 mgvfvlslatlfgstgaltgprgnshscdvtggycfpeishlwctyspyfslade 60
 QY 61 SAISPDPVDCRVTFFQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdpvdkcrvtfvqlsrhgaryptssaskaysalialieaiqknatafkgyafilktyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgnqmvnsgikfrryrykalkarkivpfirasgsdrviasaekfiqfqs 180
 QY 181 KLADPGSOPHOASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 kladpgsophoaspvinviipgsgynntldhgtctafedstlgddveanftalfapair 240

Db 181 kladpgsqhgapvlnvipegsgyntldhglctafedstlglgdveanftalfapair 240
 QY 241 ARLEADLPQVTLTDRVYVLMDCPFDVARTSDATLSPFCALFTHDEWTDYDYLQSLG 300
 Db 241 arleadlpqvtltdedvvyldmcpfdkavrtsdatslpcalfthdewtdydyqlsgl 300
 QY 301 KYGAGNPLGPAGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyggaggnplgpagvgfanelliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIETDGYASWVPPFAARAYVEMMCOQAEKEP 420
 Db 361 hdnctmsiffalglngtkplsttsvesieetdgyaswtvppfaarayvemmqcqaekp 420
 QY 421 LVRVLNDRVPLHGCVDKLGCRKRDDEGLSFARSGGNWAECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrddfveglfsarsggnwaecefa 467

RESULT 6
 AAW93382
 ID AAW93382 standard; Protein; 467 AA.
 AC AAW93382;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal consensus phytase protein mutant Q50T.
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX
 PN EP897985-A2.
 XX
 PD 24-FEB-1999.
 XX
 PF 15-JUL-1998; 98EP-0113176.
 XX
 PR 24-JUL-1997; 97EP-0112688.
 XX
 PA (HOFF) HOFFMANN LA ROCHE AG F.
 XX
 PI Lehmann M;
 XX
 DR WPI; 1999-134647/12.
 XX
 PT Preparation of a consensus protein, especially a phytase - using
 programs to compare evolutionary similarity of sequences
 XX
 PS Claim 9; Page -; 30pp; English.
 XX

This invention describes a novel process for the preparation of a consensus protein. The specific example given in the specification is that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses phytase to valuable feed additives, or variant or muten. The method is useful for improving protein properties by altering their sequence. The consensus protein and muten are useful in food, feed or pharmaceutical compositions. This sequence is a mutant phytase protein of the invention which does not appear in the specification but has been created from the protein represented in AAW93380.

Sequence 467 AA;

Query Match 98.1%; Score 2414; DB 20; Length 467;
 Best Local Similarity 98.1%; Fred. NO. 2.7e-234;
 Matches 458; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MGFWVLLSIATLFGSTGALGRGNSHSCDVTGQYQCPFEISHLWGTSPYFSLADE 60

Db 1 mgfwvllsiatlfgstgaltgrnshscdvtggyqcfpeishlwtgspyfslede 60
 QY 61 SAISPOVPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFILKTYN 120
 Db 61 saispvdpddcrvtfvqvlshrhgaryptsskskaysalieleaigknatafkgyafikty 120
 QY 121 YTLGADDLPPFENQMVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFIQFOSA 180
 Db 121 ytlgaddlppfengmnsqikfryrykalarikiwpffirassgsdrviasaekfiqfosa 180
 QY 181 KLADPGSQHQPASVNLNIIPEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 kladpgsqhgapvlnvipegsgyntldhgtctafedselgddveanftalfapair 240
 QY 241 ARLEADLPQVTLTDEVDVYVLMDCPFDVARTSDATLSPFCALFTHDEWTDYDYLQSLG 300
 Db 241 arleadlpqvtltdedvvyldmcpfdvartsdatslpcalfthdewtdydyqlsgl 300
 QY 301 KYGAGNPLGPAGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyggaggnplgpagvgfanelliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIETDGYASWVPPFAARAYVEMMCOQAEKEP 420
 Db 361 hdnctmsiffalglngtkplsttsvesieetdgyaswtvppfaarayvemmqcqaekp 420
 QY 421 LVRVLNDRVPLHGCVDKLGCRKRDDEGLSFARSGGNWAECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrddfveglfsarsggnwaecefa 467

RESULT 7
 AAW93380
 ID AAW93380 standard; Protein; 467 AA.
 AC AAW93380;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal phytase protein consensus DNA.
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX
 PN EP897985-A2.
 XX
 PD 24-FEB-1999.
 XX
 PF 15-JUL-1998; 98EP-0113176.
 XX
 PR 24-JUL-1997; 97EP-0112688.
 XX
 PA (HOFF) HOFFMANN LA ROCHE AG F.
 XX
 PI Lehmann M;
 XX
 DR WPI; 1999-134647/12.
 XX
 N-PSDB; AAX23022.
 XX
 PT Preparation of a consensus protein, especially a phytase - using
 programs to compare evolutionary similarity of sequences
 XX
 PS Claim 8; Fig 2; 30pp; English.
 XX

This invention describes a novel process for the preparation of a consensus protein. The specific example given in the specification is that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses phytase to valuable feed additives, with a fully defined amino acid sequence given in the specification, or variant or muten. The method is

CC useful for improving protein properties by altering their sequence. The
 CC consensus protein and muten are useful in food, feed or pharmaceutical
 CC compositions. This sequence represents the consensus phytase protein used
 CC in the method of the invention.

XX SQ Sequence 467 AA;

Query Match 97.8%; Score 2408; DB 20; Length 467;

Best Local Similarity 97.9%; Pred. No. 1.1e-233;
 Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCPEISHLWGTSPYFSLADE 60

DB 1 mgvfvlslatlfgstgtalgrgnshscdvtvdggycfpelshlwgtpyfsfslade 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120

DB 61 saispdvdcrvtfvqlsrhgaryptsskaysalialieaigknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKIEGQSA 180

DB 121 ytlgaddltpfgengmrvnsigkifryrykalkarkivpfirasgsdrviasaekiegiqsa 180

QY 181 KLADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240

DB 181 kladpgsqphqaspvldviiepgsgynntldhgtctafedseelgddveanftalfapair 240

QY 241 ARLEADLPQVTLTDEDVYVLMDCPFTVARTSDATELSDFCALTFHDEWIOYDYLSLG 300

DB 241 arleadlpqvtltddedvvyldmcpftvartsdatselspfcalfthdewioydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

DB 301 kygygagnplgpqagvgfaneliarlthspvqdhstntnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMOCQAEKEP 420

DB 361 hdnsmisiffalglngtkplsttsvesieetdgyaswvtpfarayvemmqcqaekp 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467

DB 421 lvrvlvndrvvplhgcavdkgcrkddfveglsfarsggnwaecfa 467

RESULT 8

AAW93381

ID AAW93381 standard; Protein; 467 AA.

XX AC

XX AAW93381;

XX AC

DT 11-JUN-1999 (first entry)

XX DE

XX Fungal consensus phytase protein mutant Q50L.

XX DE

XX phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;

XX KW

XX feed additive; variant; muten; feed; pharmaceutical.

XX KW

XX Fungi.

XX OS

XX Synthetic.

XX XX

XX EP897985-A2.

XX PN

XX 24-FEB-1999.

XX PD

XX 15-JUL-1998; 98EP-0113176.

XX XX

XX 24-JUL-1997; 97EP-0112688.

XX XX

XX (HOFF) HOFFMANN LA ROCHE AG F.

XX PA

XX PI Lehmann M;

XX XX

DR WPI; 1999-134647/12.

XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences

XX XX

PS Claim 9; Page -: 30pp; English.

XX XX

CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.

XX XX

SQ Sequence 467 AA;

Query Match 97.8%; Score 2408; DB 20; Length 467;

Best Local Similarity 97.9%; Pred. No. 1.1e-233;

Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCPEISHLWGTSPYFSLADE 60

DB 1 mgvfvlslatlfgstgtalgrgnshscdvtvdggycfpelshlwgtpyfsfslade 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120

DB 61 saispdvdcrvtfvqlsrhgaryptsskaysalialieaigknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKIEGQSA 180

DB 121 ytlgaddltpfgengmrvnsigkifryrykalkarkivpfirasgsdrviasaekiegiqsa 180

QY 181 KLADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240

DB 181 kladpgsqphqaspvldviiepgsgynntldhgtctafedseelgddveanftalfapair 240

QY 241 ARLEADLPQVTLTDEDVYVLMDCPFTVARTSDATELSDFCALTFHDEWIOYDYLSLG 300

DB 241 arleadlpqvtltddedvvyldmcpftvartsdatselspfcalfthdewioydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

DB 301 kygygagnplgpqagvgfaneliarlthspvqdhstntnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTLSIFFALGLYNGTKPLSTTSVESIEETDGYSAWTVPFARAYVEMMOCQAEKEP 420

DB 361 hdnsmisiffalglngtkplsttsvesieetdgyaswvtpfarayvemmqcqaekp 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467

DB 421 lvrvlvndrvvplhgcavdkgcrkddfveglsfarsggnwaecfa 467

RESULT 9

AAW93381

ID AAW93381 standard; Protein; 467 AA.

XX AC

XX AAW93381;

XX AC

XX 05-DEC-2000 (first entry)

XX DT

XX Consensus phytase SEQ ID NO:16.

XX DE

XX phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

CC to utilise this form of phosphate, resulting in the addition of extra
CC phosphate to the feed of such animals. Phytase enhances the nutritional
CC value of plant material without the need for adding additional phosphate
CC to the feed. The level of phosphate pollution in the environment is
CC reduced by adding phytase to animal feed, as the animal can make use of
CC the inorganic phosphate liberated from phytate phosphorus using the
CC enzyme. The phytase formulation of the invention has an improved
CC thermostability and can therefore remain stable during long-term storage
CC and can withstand feed processing methods such as extrusion, expansion
CC and pelleting. The present sequence represents a consensus sequence,
CC designated phytase-1, which was derived from the mature phytase
CC sequences from a variety of fungi (AA95944-V69556), and additionally
CC contains the Aspergillus terreus csb116.46 signal peptide at the
CC N-terminus.
XX
XX
SQ Sequence 467 AA;

Query Match 97.8%; Score 2408; DB 21; Length 467;
Best Local Similarity 97.9%; Pred. No. 1.4e-233;
Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDVTDDGGYQCFPEISHLWGTYSPIYSLADE 60
DB 1 mgvfvlslsiatlfgstgtalgrgnshscdvtddggycfpeishlwgyspyfslade 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFILKTYN 120
DB 61 saispdvddcrvtfvqlsrhgarpytssaskaysalieaiqknatafkgyafilktyn 120

QY 121 YTLGADLLTPFGENQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQGFOSA 180
DB 121 ytlgaddlltpfgenqmvnsgkifryrykalkarkivpfirasgsdrviasaekfieqfosa 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 kladpgsqphqaspvldviipsgsynntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDTVARTSDATELSPPCALFTHDEWIQDYQLQSLG 300
DB 241 arleadlpvltltdedvvyilmcpcfetvartsdatselsspccalfthdewiqdyqlslg 300

QY 301 KYYGYGAGNPLGPAQGVGFANELIARLTSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kyygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYSAKSTVPPFAARAYVEMMQCAKEP 420
DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyssakstvpfgarayvemmqcakep 420

QY 421 LVRVLVNDVRVPLHGCAVDKLGCRKRDDEFVGLSFARSGGNWAECEFA 467
DB 421 lvrvlvndrvvplhgcavdklgrckrddfvglsfarsggnwaecefa 467

RESULT 11
AAW93383
ID AAW93383 standard; Protein; 467 AA.
XX
XX AAW93383;
AC AAW93383;
XX
XX 11-JUN-1999 (first entry)
DT
XX Fungal consensus phytase protein mutant Q50G.
DE
XX Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
KW feed additive; variant; mutin; feed; pharmaceutical.
XX
XX Fungl.
OS Synthetic.
XX
XX EP897985-A2.
PN
XX

PD 24-FEB-1999.
XX
XX
PF 15-JUL-1998; 98EP-0113176.
XX
XX
PR 24-JUL-1997; 97EP-0112688.
XX
XX
PA (HOFF) HOFFMANN LA ROCHE AG F.
XX
XX Lehmann M;
PI
XX
XX WPI; 1999-134647/12.
DR
XX
XX preparation of a consensus protein, especially a phytase - using
PT programs to compare evolutionary similarity of sequences
PT
XX
XX Claim 9; Page -: 30pp; English.
PS
XX
XX This invention describes a novel process for the preparation of a
CC consensus protein. The specific example given in the specification is
CC that of a fungal phytase (myo-inositol hexakisphosphate) which
CC hydrolyses phytase to valuable feed additives, with a fully defined amino
CC acid sequence given in the specification, or variant or mutin. The
CC method is useful for improving protein properties by altering their
CC sequence. The consensus protein and mutin are useful in food, feed or
CC pharmaceutical compositions. This sequence is a mutant phytase protein of
CC the invention which does not appear in the specification but has been
CC created from the protein represented in AAW93380.
XX
XX
SQ Sequence 467 AA;

Query Match 97.8%; Score 2407; DB 20; Length 467;
Best Local Similarity 97.9%; Pred. No. 1.4e-233;
Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDVTDDGGYQCFPEISHLWGTYSPIYSLADE 60
DB 1 mgvfvlslsiatlfgstgtalgrgnshscdvtddggycfpeishlwgyspyfslade 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAPKGYAFILKTYN 120
DB 61 saispdvddcrvtfvqlsrhgarpytssaskaysalieaiqknatafkgyafilktyn 120

QY 121 YTLGADLLTPFGENQMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQGFOSA 180
DB 121 ytlgaddlltpfgenqmvnsgkifryrykalkarkivpfirasgsdrviasaekfieqfosa 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
DB 181 kladpgsqphqaspvldviipsgsynntldhgtctafedselgddveanftalfapair 240

QY 241 ARLEADLPVTLTDEDVYVLMDCPFDTVARTSDATELSPPCALFTHDEWIQDYQLQSLG 300
DB 241 arleadlpvltltdedvvyilmcpcfetvartsdatselsspccalfthdewiqdyqlslg 300

QY 301 KYYGYGAGNPLGPAQGVGFANELIARLTSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kyygygagnplgpaqgvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNMTISIFFALGLYNGTKPLSTTSVESIEETDGYSAKSTVPPFAARAYVEMMQCAKEP 420
DB 361 hdnsmisiffalglyngtkplsttsvesieetdgyssakstvpfgarayvemmqcakep 420

QY 421 LVRVLVNDVRVPLHGCAVDKLGCRKRDDEFVGLSFARSGGNWAECEFA 467
DB 421 lvrvlvndrvvplhgcavdklgrckrddfvglsfarsggnwaecefa 467

RESULT 12
AAW93384
ID AAW93384 standard; Protein; 467 AA.
XX
XX AAW93384;
AC

XX 11-JUN-1999 (first entry)
 XX Fungal consensus phytase protein mutant Q50T-Y51N.
 XX
 DE
 XX
 KW phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 PN EP897985-A2.
 XX 24-FEB-1999.
 PD 15-JUL-1998; 98EP-0113176.
 PF 24-JUL-1997; 97EP-0112688.
 XX
 PR (HOFF) HOFFMANN LA ROCHE AG F.
 XX Lehmann M;
 DR WPI; 1999-134647/12.
 XX
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 PS Claim 9; Page -: 30pp; English.
 XX
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.
 XX
 SQ Sequence 467 AA;

Query Match 97.7%; Score 2405; DB 20; Length 467;
 Best Local Similarity 97.9%; Pred. No. 2.2e-233;
 Matches 457; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCFPEISHLWGTYSPTFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshtcdtdvggyqcfpeishlwtgtnspysfslade 60
 QY 61 SAISPDPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNATAFKGYAFKLTYN 120
 Db 61 saispdpddcrvtfvqlsrhgaryptsaskaysalialeakgnatafkgyafkltyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgnqmvnsgikfryrykalkarkivpfiirasgsdrviasaekfiegfosa 180
 QY 181 KLADPGSQPHQASPVNLVITPEGSYNNLTGCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 kladpgsqphqaspvldilpegsynntldhgtctafedselgddveanftalfapair 240
 QY 241 ARLEADLPGLVTLTDEDVYVLMDCPFDTVARTSDATELSPFCALTFHDEWQYDYLQSLG 300
 Db 241 arleadlpvltltdedvylmcmcfetvartsdatselspfcalfthdewrgydyqlsg 300
 QY 301 KYGAGNPLCPAGGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyggagnplpaggvgfaneliarltrspvqdhstntnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCAKEP 420

Db 361 hdnsmisiffalgllyngtaplstsvesleedgysaswtvpfgarayvemmqcqaekp 420
 QY 421 LVRVLVNDRVVPLHGCAYDKLGRCKRDEVEGLSFARSGGNNAECFA 467
 Db 421 lvrvlvndrvvplhgcaydklgrckrddvveglisfarsggnwaecfa 467
 RESULT 13
 AAW93385
 ID AAW93385 standard; Protein; 467 AA.
 XX
 AC AAW93385;
 XX
 DT 11-JUN-1999 (first entry)
 XX
 DE Fungal consensus phytase protein mutant Q50L-Y51N.
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; muten; feed; pharmaceutical.
 XX
 OS Fungi.
 OS Synthetic.
 XX EP897985-A2.
 XX 24-FEB-1999.
 PD 15-JUL-1998; 98EP-0113176.
 PF 24-JUL-1997; 97EP-0112688.
 XX (HOFF) HOFFMANN LA ROCHE AG F.
 XX Lehmann M;
 DR WPI; 1999-134647/12.
 XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences
 PS Claim 9; Page -: 30pp; English.
 XX
 CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or muten. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and muten are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.
 XX
 SQ Sequence 467 AA;

Query Match 97.4%; Score 2399; DB 20; Length 467;
 Best Local Similarity 97.6%; Pred. No. 8.8e-233;
 Matches 456; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTGALGRGNSHSCDVTVDGGYQCFPEISHLWGTYSPTFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshtcdtdvggyqcfpeishlwtgtnspysfslade 60
 QY 61 SAISPDPDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNATAFKGYAFKLTYN 120
 Db 61 saispdpddcrvtfvqlsrhgaryptsaskaysalialeakgnatafkgyafkltyn 120
 QY 121 YTLGADDLTPFGNQMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgnqmvnsgikfryrykalkarkivpfiirasgsdrviasaekfiegfosa 180

181 KIADPGSQHQAASPVINVIIEGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240
 181 KIADPGSQHQAASPVINVIIEGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240
 241 ARLEADLPVTLTDEVDVYLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 241 ARLEADLPVTLTDEVDVYLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467
 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467

RESULT 14
 AAY39906 standard; Protein: 467 AA.

07-DEC-1999 (first entry)
 Ascomycete consensus phytase protein sequence.
 Phytase; variant; enzyme; phosphorus liberation; phytase substrate;
 Phytate level reduction; animal manure; food preparation;
 soy processing; inositol manufacture.

Synthetic.
 WO9949022-A1.
 30-SEP-1999.
 22-MAR-1999; 99WO-DK00153.
 23-MAR-1998; 98DK-0000407.
 19-JUN-1998; 98DK-0000806.
 18-SEP-1998; 98DK-0001176.
 22-JAN-1999; 99DK-0000091.

(NOVO) NOVO-NORDISK AS.
 Svendsen A;
 WPI; 1999-580444/49.
 N-PSDB; AA627423.
 New variant phytase enzymes, used for liberating phosphorus from a
 phytase substrate, for reducing phytate levels in animal manure and in
 feed and food preparations
 Claim 26; Fig 9a-d; 14lpp; English.

This sequence represents the ascomycete consensus phytase sequence.
 The invention relates to variant phytase enzymes with specific amino acid
 substitutions for improved properties. The phytase variants can be used
 for liberating phosphorus from a phytase substrate. They can be used for
 reducing phytate levels in animal manure. They can be used in feed or
 food preparations. The phytase DNA can also be used to produce transgenic
 plants which can be used in feeds or foods. The phytase variants can also
 be used in soy processing and in the manufacture of inositol or
 derivatives. The phytase variants can have altered activities such as pH
 stability, temperature stability, pH profile, temperature profile,
 specific activity (in particular in relation to pH and temperature),
 substrate specificity, substrate cleavage pattern, substrate binding,

CC position specificity, the velocity and level of release of phosphate
 CC from corn, reaction rate, phytase degradation rate and end level of
 CC released phosphate reached.
 XX
 SQ Sequence 467 AA;
 Query Match 97.4%; Score 2398; DB 20; Length 467;
 Best Local Similarity 97.4%; Pred No. 1,le-232;
 Matches 455; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MGFFVLLSLATLFGSTGALGPRGNHSCDTVDGGYOCFPEISHLMCTGTYSPYSLADE 60
 Db 1 mgffvllslatlfstgaglylgrgnshscdtvdggyocfpeishlmctgtyspyslede 60
 QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFKTYN 120
 Db 61 saispdvddcrvtfvqvlsrhgaryptssaskaysalieleaigknatafkgyafikty 120
 QY 121 YTLGADDLTPFGENOMVNSGIRFYRYKALARKIVPFIIRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgenomvnsigirfyrykalarakiyfirasgsdrviasaekfiqfqs 180
 QY 181 KIADPGSQHQAASPVINVIIEGSGYNNLTLDHGCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 kiadpgsqhqaaspviviepgsgynntldhgctafedselgddveanftalfapair 240
 QY 241 ARLEADLPVTLTDEVDVYLMDCMPCFDTVARTSDATLSLPCALFTHDEWIOYDYLQSLG 300
 Db 241 arleadlpvltldevdvyilmcmptetvartsdatselcpalfthdewryqylqslg 300
 QY 301 KYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kyygagngplgpaqgvgfaneliarltrspvqdhstntldsnpatfplnatlyadfs 360
 QY 361 HDNTMISIFFALGLYNGTKPLSTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 Db 361 hdnsmisiffalglyngtkplstsvesietdgyaswtvpfagarayvemmqcqaekp 420
 QY 421 LVRVLVNDVRVPLHGCVDKLGRCRDKDFVEGLSFARSGGNWAECEFA 467
 Db 421 lvrvlvndrvvplhgcavdklgrckrddfveglsfarsggnwaecefa 467

RESULT 15
 AAY43170
 ID AAY43170 standard; Protein: 467 AA.
 XX
 AC AAY43170;
 XX
 DT 06-JAN-2000 (first entry)
 XX
 DE Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.
 XX
 KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 XX
 OS consensus sequence.
 XX
 OS Synthetic.
 XX
 PN WO994380-A1.
 PD 30-SEP-1999.
 XX
 PF 22-MAR-1999; 99WO-DK00154.
 XX
 PR 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.
 PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.
 PR 22-JAN-1999; 99DK-0000093.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX

```

PI Petersen S;
XX
DR WPI; 1999-591030/50.
DR N-PSDB; AAZ31521.
XX
PT Preparing animal feed using a thermostable phytase
XX
PS Example 3; Fig 10; 7lpp; English.
XX
CC This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
CC The invention relates to a process for preparing animal feed by
CC agglomerating feed ingredients with a thermostable phytase, which is
CC added before or during agglomeration. The thermostable phytase is useful
CC for expression in transgenic plants. These plants are useful in the
CC preparation of animal feed itself. The thermostable phytase allows animal
CC feed to be produced more efficiently, in addition to improved
CC phytase-expressing transgenic plants. These plants provide a feed
CC ingredient and a feed additive (phytase) simultaneously.
XX
SQ Sequence 467 AA;

Query Match 95.9%; Score 2362; DB 20; Length 467;
Best Local Similarity 95.5%; Pred. No. 4.7e-229;
Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFWLLSIATLFGSTGTALGPRGNHSCDVTVDGGYQCFPEISHLWGTYSYFSLADE 60
Db 1 mgfwllsiatlfgstgtalgrgnshscdvtvdggycfpeishlwtysyffslade 60

QY 61 SAISPDVPDCRVTFFQVLSRHGARYPTSSASKAYSALIEAIOKNATAPKGYAFLKTYN 120
Db 61 saispdvpcrvtffqvlshrhgaryptssaskaysalieaioknatarkgyafltktn 120

QY 121 YTLGADDLTPFGENQWNSGIKFYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
Db 121 ytlgaddltpfgenqwnsgikfyrkyalarkivpfirasgsdrviasaekfieqfqa 180

QY 181 KLADPGSOPHOASPVINVIPEGSGYNNLTDHGCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 kladpganphqaspvinviipegsgynntldhglctafeeselqddveanftavfappir 240

QY 241 ARLEADLPGYTLTDEDVYVLMDCPFDTVARTSDATLSPFCALFTHDEWIOYDYLOSLG 300
Db 241 arleahlpgvnltdedvnlmdmcpfdtvartsdacqlspfcldfthdewigydylqsig 300

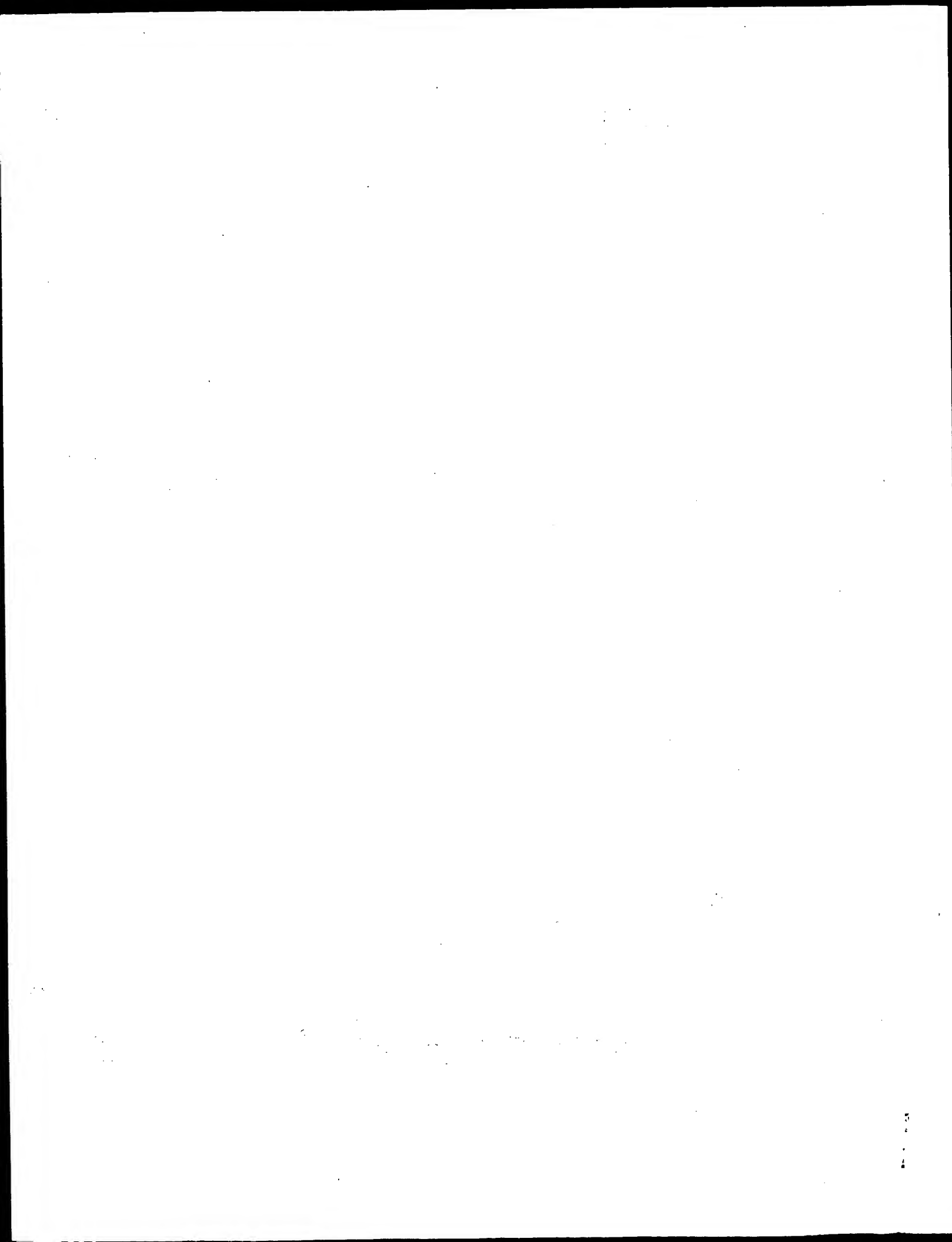
QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
Db 301 kygygagnplgpaqgvgvneliarlthspvqdhstnhtldsnpatfplnatlyads 360

QY 361 HDNTMISIFPALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
Db 361 hdnmtisiffalglyngtkplsttsvesieetdgyaswtvpfaarayvemmqqcaakep 420

QY 421 LVRVLNDRVPLHGCAVDKLGKCRKDDFVGLSFARSGGNNAECFA 467
Db 421 lvrvlndrvplhgcvdklgrckrddfvglsfarsggnweecfa 467

```

Search completed: October 26, 2001, 16:40:04
Job time: 4955 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:46 ; Search time 32.57 seconds
(without alignments)
295.232 Million cell updates/sec

Title: US-09-488-265-29
Perfect score: 2462
Sequence: 1 MGFFVLLSLATLFGSTGCT.....DFVEGLSPARSGNNAECFA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2282	92.7	441	4	US-09-121-425-1
2	2270	92.2	467	4	US-09-121-425-2
3	1879	76.3	467	1	US-07-923-724-8
4	1879	76.3	467	2	US-08-609-426A-8
5	1879	76.3	467	2	US-08-374-652C-2
6	1875	76.2	467	1	US-08-151-574-32
7	1875	76.2	467	1	US-08-146-424-20
8	1875	76.2	467	1	US-08-693-709-2
9	1875	76.2	467	2	US-08-419-448-32
10	1875	76.2	467	2	US-08-819-825-3
11	1875	76.2	467	2	US-08-163-642-3
12	1860	75.5	467	4	US-09-155-855-3
13	1818	73.8	443	4	US-09-155-855-1
14	1813	73.6	443	4	US-09-155-855-2
15	1334	54.2	475	2	US-08-819-825-2
16	1334	54.2	475	4	US-09-163-642-2
17	802.5	32.6	443	3	US-08-993-359-30
18	798.5	32.4	439	3	US-08-993-359-24
19	798.5	32.4	439	3	US-09-221-654-2
20	798.5	32.4	439	3	US-08-989-358A-2
21	795.5	32.3	453	3	US-08-993-359-22
22	781	31.7	442	3	US-08-993-359-28
23	778	31.6	442	3	US-08-993-359-26
24	363	14.7	468	1	US-07-627-539G-2
25	358.5	14.6	446	1	US-07-627-539G-7
26	347.5	14.1	479	1	US-07-923-724-2
27	347.5	14.1	479	2	US-08-609-426A-2

28	347.5	14.1	479	2	US-08-374-652C-4	Sequence 4, Appli
29	170.5	6.9	449	3	US-08-680-506-7	Sequence 7, Appli
30	148.5	6.0	92	3	US-08-993-359-32	Sequence 32, Appli
31	122	5.0	113	1	US-08-241-853-8	Sequence 8, Appli
32	122	5.0	113	2	US-08-850-917-8	Sequence 8, Appli
33	119	4.8	24	2	US-08-374-652C-32	Sequence 32, Appli
34	118	4.8	318	3	US-08-680-506-3	Sequence 3, Appli
35	111	4.5	113	1	US-08-241-853-10	Sequence 10, Appli
36	111	4.5	113	2	US-08-850-917-10	Sequence 10, Appli
37	96.5	3.9	386	1	US-08-758-213-1	Sequence 1, Appli
38	96.5	3.9	386	2	US-08-692-787-48	Sequence 48, Appli
39	96.5	3.9	386	4	US-09-097-199-48	Sequence 48, Appli
40	96.5	3.9	515	2	US-09-146-283-2	Sequence 2, Appli
41	96.5	3.9	515	3	US-08-579-823A-2	Sequence 2, Appli
42	95.5	3.9	515	4	US-09-344-195-2	Sequence 2, Appli
43	95.5	3.9	1732	2	US-08-570-311-10	Sequence 10, Appli
44	95.5	3.9	1732	2	US-08-353-485-10	Sequence 10, Appli
45	93.5	3.8	20	1	US-07-923-724-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1
US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-1

Query Match	92.7%	Score	2282	DB	4	Length	441
Best Local Similarity	97.7%	Pred. No.	1.1e-230				
Matches	431	Conservative	3	Mismatches	7	Indels	0
						Gaps	0
QY	27	NSHSCDVTGGYQCPEISHLWGTSYPSFSLADESAISPDVDDCRVTFVQVLSRHGARY	86				
Db	1	NSHSCDVTGGYQCPEISHLWGTSYPSFSLADESAISPDVDDCRVTFVQVLSRHGARY	60				
QY	87	PTSSASKAYSALIEAIQKNATAFKGYAFLKTYNTLCADDLTPGENOMVNSGKFKYRR	146				
Db	61	PTSSASKAYSALIEAIQKNATAFKGYAFLKTYNTLCADDLTPGENOMVNSGKFKYRR	120				
QY	147	YKALARKIVPFRASGSDRVATSAEKFTEGQSACLADPGSOPHOASPVINVIPEGSGY	206				
Db	121	YKALARKIVPFRASGSDRVATSAEKFTEGQSACLADPGSOPHOASPVINVIPEGSGY	180				
QY	207	NNTLDHGCTAFDESELGDDVEANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCPF	266				
Db	181	NNTLDHGCTAFDESELGDDVEANFTALFAPAIRARLEADLPVTLTDEDVYVLMDCPF	240				
QY	267	DTVARTSDATSLSPFCALFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFANELIAR	326				
Db	241	ETVARTSDATSLSPFCALFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFANELIAR	300				
QY	327	LTHSPVQDHTSTNHTLDSNPATPLNATLYADFSDHNTWISIFFALGLYNGTKPLSTTSV	386				


```
Db 301 LTRSPQDHTSTNHTLDSNPATFPLNATLYADFSDHNSWISIFFALGLYNGTAPLSTTSV 360
Qy 387 ESIEETDGYASWTVPFAARAYVEMMOQAEKEPLRVLVNDRVPLHGCADVKGRCR 446
Db 361 ESIEETDGYASWTVPFAARAYVEMMOQAEKEPLRVLVNDRVPLHGCADVKGRCR 420
Qy 447 DDFVEGLSFARSGGNWAECEFA 467
Db 421 DDFVEGLSFARSGGNWAECEFA 441

RESULT 2
US-09-121-425-2
; Sequence 2, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2

Query Match 92.2%; Score 2270; DB 4; Length 467;
Best Local Similarity: 89.7%; Pred. No. 2,1e-229;
Matches 437; Conservative 3; Mismatches 7; Indels 40; Gaps 2;

Qy 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGGYQCFFPEISHLWGTSPYFSLADE 60
Db 1 MGVFVLLSTATLFGSTGALPGRNHSCDVTGGYQCFFPEISHLWGTSPYFSLADE 60
Qy 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYSAKALIEAIQKNATAFKGYAFLKTN 120
Db 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYSAKALIEAIQKNATAFKGYAFLKTN 100
Qy 121 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
Db 101 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 160
Qy 181 KLADPGSQPHQASPVIN-----VITPEGSYNNITLDHGTCTAFED 220
Db 161 KLADPGSQPHQASPVINDLIEAIQKNATAFKGYAFLKVIIEPGSGYNNITLDHGTCTAFED 220
Qy 221 SELGDDVEANFTALFAIRLEADLPVTLDEDDVYVLMDCPDTVARTSDATELSP 280
Db 221 SELGDDVEANFTALFAIRLEADLPVTLDEDDVYVLMDCPDTVARTSDATELSP 280
Qy 281 FCALFTHDEWIQDYLSGLKYYGYGAGNPLGPAQGVGFANELIARLTHTSPVODHTSTNH 340
Db 281 FCALFTHDEWRQDYLSGLKYYGYGAGNPLGPAQGVGFANELIARLTHTSPVODHTSTNH 340
Qy 341 TLDSPATFPLNATLYADFSDHNTMISIFFALGLYNGTAPLSTTSVESIEETDGYASWT 400
Db 341 TLDSPATFPLNATLYADFSDHNTMISIFFALGLYNGTAPLSTTSVESIEETDGYASWT 400
Qy 401 VPFAARAYVEMMOQAEKEPLRVLVNDRVPLHGCADVKGRCRDDFVEGLSFARSGG 460
Db 401 VPFAARAYVEMMOQAEKEPLRVLVNDRVPLHGCADVKGRCRDDFVEGLSFARSGG 460
Qy 461 NWAECEFA 467
Db 461 NWAECEFA 467
```

```
Db 461 NWAECEFA 467

RESULT 3
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerström, Richard B.
; TITLE OF INVENTION: Production of phytase Degrading Enzymes
; TITLE OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/044,077
; FILING DATE: 29-APR-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8

Query Match 76.3%; Score 1879; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 2,1e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

Qy 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGGYQCFFPEISHLWGTSPYFSLADE 60
Db 1 MGVSALLPLLLVLLAGLAVPASRNOSTCDTVDQGGQCFSETHLWGVAPFESLANE 60
Qy 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYSAKALIEAIQKNATAFKGYAFLKTN 120
Db 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYSAKALIEAIQKNATAFKGYAFLKTN 120
```

QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGQSA 180
 Db 121 YSLGADDLTPGEGELVNSGKIFRYVESLRTNIIPIFIRSGSRVIAASGEKFTGQST 180
 QY 181 KLADPGSQHOASPVINVIIEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 KLKDPRAQPGQSSPKIDVIVISEASSNNTLDPGCTCFEDELADTVEANFTAFAPAIR 240
 QY 241 ARLEADLPGLVTLTDEDVYVLMDCMCPDFTVARTSDATELSPFCALFTHDEWIQDYLOSGL 300
 Db 241 QRLNDLSGVTLTDEVTYVLMDCMSFDTISTVDTKLSPCDLFTHDEWIHYDYLQSLK 300
 QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGHGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIEBETDGYASWTVPFAARAYVEMMQCAKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAKEP 420
 QY 421 LVRVLNDRVPLHGCAYDKLGRCKRDDFVEGLSFARSGGNWAECSA 467
 Db 421 LVRVLNDRVPLHGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 4

US-08-609-426A-8
 ; Sequence 8, Application US/08609426A
 ; Patent No. 5830733
 ; GENERAL INFORMATION:
 ; APPLICANT: Nevalainen, Helena K.M.
 ; APPLICANT: Paloheimo, Marja T.
 ; APPLICANT: Miettinen-Olmonen, Arja S.K.
 ; APPLICANT: Torkkeli, Tuula K.
 ; APPLICANT: Cantrell, Michael
 ; APPLICANT: Piddington, Christopher S.
 ; APPLICANT: Rambosek, John A.
 ; APPLICANT: Turunen, Marja K.
 ; APPLICANT: Fagerstr m, Richard B.
 ; APPLICANT: Houston, Christine S.
 ; TITLE OF INVENTION: Production of phytase Degrading Enzymes
 ; TITLE OF INVENTION: in Trichoderma
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/609,426A
 ; FILING DATE: 01-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/923,724
 ; FILING DATE: 31-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/496,155
 ; FILING DATE: 19-MAR-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/044,077
 ; FILING DATE: 29-APR-1987
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: UK 8610600
 ; FILING DATE: 30-APR-1986
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Reed, Grant E.

; REGISTRATION NUMBER: P-41,264
 ; REFERENCE/DOCKET NUMBER: 1050.0080001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-609-426A-8

Query Match 76.3%; Score 1879; DB 2; Length 467;

Best Local Similarity 75.6%; Pred. No. 2.1e-188;

Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVVLISATLFGSTGTALGPRGNHSCDTPDGGYQCFPEISHLMGTSPYFSLADE 60
 Db 1 MGVSALLPLLYLLAGVTSGLAVPASRNQSTCDTPDGGYQCFSETSHLWGOYAPFSLANE 60
 QY 61 SAISDPDPCDRCVTFQVLSRHGARYPTSSASKAYSALIEIAIKNATAFKGYAFLKTYN 120
 Db 61 SAISDPDPCDRCVTFQVLSRHGARYPTSSASKAYSALIEIAIKNATAFKGYAFLKTYN 120
 QY 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTGQSA 180
 Db 121 YSLGADDLTPGEGELVNSGKIFRYVESLRTNIIPIFIRSGSRVIAASGEKFTGQST 180
 QY 181 KLADPGSQHOASPVINVIIEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 Db 181 KLKDPRAQPGQSSPKIDVIVISEASSNNTLDPGCTCFEDELADTVEANFTAFAPAIR 240
 QY 241 ARLEADLPGLVTLTDEDVYVLMDCMCPDFTVARTSDATELSPFCALFTHDEWIQDYLOSGL 300
 Db 241 QRLNDLSGVTLTDEVTYVLMDCMSFDTISTVDTKLSPCDLFTHDEWIHYDYLQSLK 300
 QY 301 KYGYGAGNPLGPAQGVGFANELLARLTHSPVODHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGHGAGNPLGPTQGVGYANELLARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIEBETDGYASWTVPFAARAYVEMMQCAKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAKEP 420
 QY 421 LVRVLNDRVPLHGCAYDKLGRCKRDDFVEGLSFARSGGNWAECSA 467
 Db 421 LVRVLNDRVPLHGCPIDALGRCTRDSFVRGLSFARSGGDWAECSA 467

RESULT 5

US-08-374-652C-2
 ; Sequence 2, Application US/08374652C
 ; Patent No. 5834286
 ; GENERAL INFORMATION:
 ; APPLICANT: NEVALAINEN, HELENA K.M.
 ; APPLICANT: PALOHEIMO, MARJA T.
 ; APPLICANT: FAGERSTROM, RICHARD B.
 ; APPLICANT: MIETTINEN-OINONEN, ARJA S.
 ; APPLICANT: TURUNEN, MARJA K.
 ; APPLICANT: RAMBOSEK, JOHN A.
 ; APPLICANT: PIDDINGTON, CHRISTOPHER S.
 ; APPLICANT: HOUSTON, CHRISTINE S.
 ; APPLICANT: CANTRELL, MICHAEL A.
 ; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
 ; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
 ; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
 ; CITY: WASHINGTON

```

; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,652C
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07058
; FILING DATE: 27-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/925,401
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, GRANT E.
; REGISTRATION NUMBER: 41,264
; REFERENCE/DOCKET NUMBER: 1050.071001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-374-652C-2

```

```

Query Match 76.3%; Score 1879; DB 2; Length 467;
Best Local Similarity 75.6%; Pred. No. 2.1e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVEVLLSLATLFGSTGTALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPIYFSLADE 60
DB 1 MGVSALLPLLYLLAGTSLGLAVPASRNQSCDVTVDGGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAFKGYAFKLTYN 120
DB 61 SAISDPVAGCRVTFQAQVLSRHGARYPTESKGYKYSALIEEQNVTFDFGKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQWNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFRYQYESTLNRILPFISSGSSRVIASGEKFIQFQST 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTAFAPAIR 240
DB 181 KLKDPRAQPGQSSPKIDVVIASESSNNLTDPGCTVFEDELADTVEANFTAFAPAIR 240

QY 241 ARLEADLPGVTITDEQVYVLMDCPFDVARTSDATELSPPFCALFTHDEWIOYDYLQSLG 300
DB 241 ORLENDLSGVTITDEVTYVLMDCSFDTISTSTVDTKLSPPFCDLFTHDEWIHYDYLQSLK 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNAILYADF 360
DB 301 KYGYGAGNPLGTPQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADF 360

QY 361 HONTMITSIFALGLYNGTKPLSTTSVESTTEETDGSASKTVPFAARAVEMQCAQEP 420
DB 361 HONGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMQCAQEP 420

QY 421 LVRVLVNDRVPLHGCVDKLGCRKRDDEVEGLSFARSGNWAECFA 467
DB 421 LVRVLVNDRVPLHGCPIDALGCRTRDSEVRLSARSGGDWAECFA 467

```

```

RESULT 6
US-08-151-574-32
; Sequence 32, Application US/08151574
; Patent No. 5436156
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorcom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selden
; TITLE OF INVENTION: Cloning and Expression of Microbial
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025-3471
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,574
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/688,578
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-151-574-32

```

```

Query Match 76.2%; Score 1875; DB 1; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVEVLLSLATLFGSTGTALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPIYFSLADE 60
DB 1 MGVSALLPLLYLLAGTSLGLAVPASRNQSCDVTVDGGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISDPVDDCRVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAFKGYAFKLTYN 120
DB 61 SVISPEVPACRVTFQAQVLSRHGARYPTDSKGYKYSALIEEQNVTFDFGKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQWNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGQELVNSGKIFRYQYESTLNRILPFISSGSSRVIASGEKFIQFQST 180

QY 181 KLADPGSQPHQASPVINVIIPGSGYNNLTLDHGTCTAFEDSELGDDVEANFTAFAPAIR 240
DB 181 KLKDPRAQPGQSSPKIDVVIASESSNNLTDPGCTVFEDELADTVEANFTAFAPAIR 240

QY 241 ARLEADLPGVTITDEQVYVLMDCPFDVARTSDATELSPPFCALFTHDEWIOYDYLQSLG 300
DB 241 ORLENDLSGVTITDEVTYVLMDCSFDTISTSTVDTKLSPPFCDLFTHDEWIHYDYLQSLK 300

```

QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNNHTLDSNPATFPPLNATLYADFS 360
 DB 301 KYYGAGNPLGPTQGGVYANELIARLTHSPVHDDTSSNHTLDSNPATFPPLNATLYADFS 360
 QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIETDGYASASWTVPFAARAYVEMMQCAKEP 420
 DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQTDGFSAWTVPFASRLYVEMMQCAKEP 420
 QY 421 LVRVLNDRVPLHGCVAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLNDRVPLHGCVPDVALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 7

US-08-146-424-20
 ; Sequence 20, Application US/08146424
 ; Patent No. 5593963
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OOLJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/146,424
 FILING DATE: 02-NOV-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KENNEDY, BILL
 REGISTRATION NUMBER: 33,407
 REFERENCE/DOCKET NUMBER: 44615-20011.24
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-146-424-20

Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 5.6e-186;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALCPNGNSHSCDTPDGGYQCFPEISHLWGTSPYFSLADE 60
 DB 1 MGVSALLPLLYLGSVTGLAVPASRNQSSCDTPDGGYQCFSETSHLWGQAPFFSLANE 60
 QY 61 SAISFDVDDCRVTFVQLSRHGARYPTSSASKAYSALIEAIQKNATAFKGKAFKTYN 120
 DB 61 SVISPEVAGCRVTFQAQLSRHGARYPTDSKGYKYSALIEEQQNTATFDCKYAFKTYN 120
 QY 121 YTLGADLLTPFENQWNSGKIFYRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180

DB 121 YSLGADLLTPFGEQELVNSGKIFYRYESLTHSNVFFIRSSGSSRVIASGKFIQFQST 180
 QY 181 KLADPGSQHQSPVNTVPIPEGSGYNNLTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 KLKDPRAQPGQSPKIDVVISSEASSNNLTDPGTCTVFEDSELADTVEANFTATVPISR 240
 QY 241 ARLEADLPGVTLTDEDDVYVLMDCPFDIVARTSDATELSPPFCALTHDEWIOYDYQLSLG 300
 DB 241 QRLNDLSGVTLTDEVTYVLMDCSPFDITSTVDTKLSPPFCDLTHDEWINYDYQLSLK 300
 QY 301 KYYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNNHTLDSNPATFPPLNATLYADFS 360
 DB 301 KYYGAGNPLGPTQGGVYANELIARLTHSPVHDDTSSNHTLDSNPATFPPLNATLYADFS 360
 QY 361 HDNTWISIFFALGLYNGTKPLSTTSVESIETDGYASASWTVPFAARAYVEMMQCAKEP 420
 DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQTDGFSAWTVPFASRLYVEMMQCAKEP 420
 QY 421 LVRVLNDRVPLHGCVAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLNDRVPLHGCVPDVALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 8

US-08-693-709-2
 ; Sequence 2, Application US/08693709
 ; Patent No. 5770413
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN OOLJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
 ; TITLE OF INVENTION: IN PLANTS
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/693,709
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146,424
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20011.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-813-5600
 ; TELEFAX: 415-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal

FEATURE:
 NAME/KEY: Signal Sequence
 LOCATION: 1...23
 OTHER INFORMATION:
 US-08-693-709-2

Query Match 76.2%; Score 1875; DB 1; Length 467;
 Best Local Similarity 75.6%; Pred. No. 5.6e-188;
 Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTVDGGYOCPEISHLWCTYSPYFSLADE 60
 DB 1 MGVSAVLLPLLYLSGVTSGLVAPASRNQSCDVTVDGGYOCPEISHLWCTYSPYFSLANE 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSAIEAIQKNATAFKGYAFLKTYN 120
 DB 61 SVISPEVPACGRVTFQAQLSRHGARYPTDSKGYKSALIEEQONATTFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFEIGFQSA 180
 DB 121 YSLGADDLTPFGELVNSGKIFRYRYKALARKIVPFFIRSSGSSRVIASGKKFIEGFQST 180

QY 181 KLADPGSOPHQAASPVNIVLPEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 KLKDPRAQCGSPKIDVIVSEASSNNTLDPCCTCTVEDSELADTVEANFTATFVPSIR 240

QY 241 ARLEADLPVTLTDEVDVYLMDCPFDVARTSDATELSPCALFTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPCDFLTHDEWINDYQLQSLK 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 DB 301 KYGHGAGNPLGPTQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 9

US-08-419-448-32
 Sequence 32, Application US/08419448
 Patent No. 5863533

GENERAL INFORMATION:

APPLICANT: Robert F.M. Van Gorcom
 APPLICANT: Willem Van Hartingsveldt
 APPLICANT: Petrus A. Van Paridon
 APPLICANT: Annemarie E. Veenstra
 APPLICANT: Rudolf G.M. Luttin
 APPLICANT: Gerardus Seiten
 TITLE OF INVENTION: Cloning and Expression of Microbial
 TITLE OF INVENTION: Phytase
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 SOFTWARE:
 CURRENT APPLICATION DATA: US/08/419,448
 APPLICATION NUMBER: US/08/419,448

FILING DATE: 10-APR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 24615-20026.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-887-1500
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 467 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-419-448-32

Query Match 76.2%; Score 1875; DB 2; Length 467;

Best Local Similarity 75.6%; Pred. No. 5.6e-188;

Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTVDGGYOCPEISHLWCTYSPYFSLADE 60
 DB 1 MGVSAVLLPLLYLSGVTSGLVAPASRNQSCDVTVDGGYOCPEISHLWCTYSPYFSLANE 60

QY 61 SAISPDVDDCRVTFVQVLSRHGARYPTSSASKAYSAIEAIQKNATAFKGYAFLKTYN 120
 DB 61 SVISPEVPACGRVTFQAQLSRHGARYPTDSKGYKSALIEEQONATTFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGENQVNSGKIFRYRYKALARKIVPFFIRASGSDRVIASAEKFEIGFQSA 180
 DB 121 YSLGADDLTPFGELVNSGKIFRYRYKALARKIVPFFIRSSGSSRVIASGKKFIEGFQST 180

QY 181 KLADPGSOPHQAASPVNIVLPEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
 DB 181 KLKDPRAQCGSPKIDVIVSEASSNNTLDPCCTCTVEDSELADTVEANFTATFVPSIR 240

QY 241 ARLEADLPVTLTDEVDVYLMDCPFDVARTSDATELSPCALFTHDEWIQDYQLQSLG 300
 DB 241 QRLNDLSGVTLDTEVTYLMDCSFDTISTSTVDTKLSPCDFLTHDEWINDYQLQSLK 300

QY 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVODHTSTNHTLDSNPATFPLNATLYADF 360
 DB 301 KYGHGAGNPLGPTQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADF 360

QY 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 DB 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSAWTVPFASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
 DB 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 10

US-08-819-825-3

Sequence 3, Application US/08819825

Patent No. 5866118

GENERAL INFORMATION:

APPLICANT: Berk, Randy M.
 APPLICANT: Ray, Michael W.
 APPLICANT: Klotz, Alan V.
 TITLE OF INVENTION: Polypeptides Having Phytase Activity
 TITLE OF INVENTION: And Nucleic Acids Encoding Same
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5866118 of No. 5866118 disk of No. 5866118th America, Inc.
 STREET: 405 Lexington Avenue, Suite 6400

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10174-6401

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-819-825-3

```

```

Query Match          76.2%; Score 1875; DB 2; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

```

```

Qy 1 MGVEFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTSPYFSLADE 60
Db 1 MGVSALLPLXLLSGVTSGLAVPASRNQSCDVTVDGGYQCFSETSHLWGOYAPFSLANE 60

Qy 61 SATSPDVPDDCRVTFVVOVLSRHGARYPTSSAKYASALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFQAQVLSRHGARYPTDSKGKYSALIEIEIQONATTEDGKYAFLKTYN 120

Qy 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTIEGQSA 180
Db 121 YSLGADDLTPFGEOLVNSGKIFQRYESLTRNIVPFISSGSSRVIASGKFTIEGFQST 180

Qy 181 KLADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLADPRAQPGQSSPKIDVISEASSNNTLDPGCTVFEDELADTVEANFTATFVPSIR 240

Qy 241 ARLEADLPGLVTLDEDDVYLMDCMCPDTVARTSDATFELSPFCALFTHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVTYLMDCMCSFTISTSTVDTKLSPFCOLFTHDEWIQDYLSLK 300

Qy 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGTPQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360

Qy 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASATVTPFAARAYVEMMOCAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSATVTPFASRLYVEMMOCAQEP 420

Qy 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
Db 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSEFVRLGSFARSGGDWAECEFA 467

```

```

RESULT 11
US-09-163-642-3
; Sequence 3, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: No. 6221644o No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-163-642-3

```

```

Query Match          76.2%; Score 1875; DB 4; Length 467;
Best Local Similarity 75.6%; Pred. No. 5.6e-188;
Matches 353; Conservative 41; Mismatches 73; Indels 0; Gaps 0;

```

```

Qy 1 MGVEFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTSPYFSLADE 60
Db 1 MGVSALLPLXLLSGVTSGLAVPASRNQSCDVTVDGGYQCFSETSHLWGOYAPFSLANE 60

Qy 61 SATSPDVPDDCRVTFVVOVLSRHGARYPTSSAKYASALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFQAQVLSRHGARYPTDSKGKYSALIEIEIQONATTEDGKYAFLKTYN 120

Qy 121 YTLGADDLTPFGENOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTIEGQSA 180
Db 121 YSLGADDLTPFGEOLVNSGKIFQRYESLTRNIVPFISSGSSRVIASGKFTIEGFQST 180

Qy 181 KLADPGSQPHQASPVINVIIEGSGYNNTLDHGTCTAFEDSELGDDVEANFTALFAPAIR 240
Db 181 KLADPRAQPGQSSPKIDVISEASSNNTLDPGCTVFEDELADTVEANFTATFVPSIR 240

Qy 241 ARLEADLPGLVTLDEDDVYLMDCMCPDTVARTSDATFELSPFCALFTHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVTYLMDCMCSFTISTSTVDTKLSPFCOLFTHDEWIQDYLSLK 300

Qy 301 KYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGYGAGNPLGTPQGVGFANELIARLTHSPVHDDTSSNHTLDSNPATFPLNSTLYADFS 360

Qy 361 HDNTMISIFFALGLYNGTKPLSTTSVESIEETDGYASATVTPFAARAYVEMMOCAKEP 420
Db 361 HDNGIISILFALGLYNGTKPLSTTTVENITQDGFSSATVTPFASRLYVEMMOCAQEP 420

Qy 421 LVRVLVNDRVVPLHGCAVDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
Db 421 LVRVLVNDRVVPLHGCPVDALGCRTRDSEFVRLGSFARSGGDWAECEFA 467

```

```

RESULT 12
US-09-155-855-3

```

; Sequence 3, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; SOFTWARE: PatentIn Ver. 2.0
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 3
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-3

Query Match 75.5%; Score 1860; DB 4; Length 467;
Best Local Similarity 74.1%; Pred. No. 2.1e-186;
Matches 346; Conservative 48; Mismatches 73; Indels 0; Gaps 0;
QY 1 MGVFVLLSIAFLFGSTGCTALCPRGNSHSDTVDGGYOCFPEISHLWGTSPYFSLADE 60
DB 1 MGVSAVLLPLXLLSGVTSGLAVPASRNGOSTCDTVDGGYOCFSETSHLWGVAPFSLANK 60
QY 61 SAISDPVDDCRVTFVQLSRHGARYPTSSAKYSAKYSALIEIAOKNATAFKGYAFLKTYN 120
DB 61 SAISDPVAGCHVTFAQLSRHGARYPTDSKGYKYSALIEIQONATTFEKGYAFLKTYN 120
QY 121 YTLGADDLTPFCENOMVNSGKIFRYRYKALARKIVPFTIRASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFCQELVNSGKVFQRYESLRNIVPFTIRSGSRVSIASGNKFIEGQST 180
QY 181 KLADGSPQHPASPVINVIIEGSGYNTLDHGCTAFEDSELGDVDFANFTALFAPAIR 240
DB 181 KLKDPRAQPCQSPKIDVVISSEASTSNNTLDPGCTCTFEDSELADDIENATFATFVPSIR 240
QY 241 ARLEADLPVTLTDEEVYLMDCFDVARTSDATELSPCALFTHDHWIOYDYLQSLG 300
DB 241 ORLENDLSGVSILTDTEVYLMDCSFDISTSTVDTKLSFPCDLTFHEEWINDYQLSLN 300
QY 301 KYGYGAGNPLGPAQCGVGFANELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYGHGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADFS 360
QY 361 HONTMISIFALGLYNGTKPLSTTSVESIEETDGYASWTFVFAARAYVEMMQCAKEP 420
DB 361 HDNGIISILFALGLYNGTKPLSTTAENITQDGFSSAWTFVFAARAYVEMMQCQSEQEP 420
QY 421 LVRVLNDRVPLHGCANDKLGCRKDDFVEGLSFARSGGNWAECEFA 467
DB 421 LVRVLNDRVPLHGPCVDALGCRTRDSEFVKGLSFARSGGDWGECEFA 467

RESULT 13
US-09-155-855-1
; Sequence 1, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Aspergillus niger
US-09-155-855-1

Query Match 73.8%; Score 1818; DB 4; Length 443;
Best Local Similarity 75.7%; Pred. No. 4.8e-182;
Matches 334; Conservative 46; Mismatches 61; Indels 0; Gaps 0;
QY 27 NNSHCDTVDGGYOCFPEISHLWGTSPYFSLADESAISDPVDDCRVTFVQVLSRHGARY 86
DB 3 NQSTCDTVDGGYOCFSETSHLWGTAPFSLANKSAISDPVAGCHVTFAQVLSRHGARY 62
QY 87 PTSSAKYSAKYSALIEIAOKNATAFKGYAFLKTYNYTLGADDLTPFCENOMVNSGKIFYRR 146
DB 63 PTDSKGYKYSALIEIQONATTFEKGYAFLKTYNSLGADDLTPFCQELVNSGKVFYOR 122
QY 147 YKALARKIVPFTIRASGSDRVIASAEKFIQFQSAKLADPGSOPHOASPVINVIIEGSGY 206
DB 123 YESLTRNIVPFTIRSGSRVSIASGNKFIEGPOSTKLKDPRAQPGSSPKIDVVISSEASTS 182
QY 207 NNTLDGCTCTAFEDSELGDVDFANFTALFAPAIRARLEADLPVTLTDEEVYLMDCPF 266
DB 183 NNTLDGCTCTFEDSELADDIENATFATFVPSIRQRLNDSVSLTDTEVYLMDCSF 242
QY 267 DTVARTSDATELSPCALFTHDHWIOYDYLQSLGKYYGACNPLGPAQCGVGFANELIAR 326
DB 243 DTISTSTVDTKLSFPCDLTFHEEWINDYQLSLNKYYGHGAGNPLGPTQGVGYANELIAR 302
QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSHONTMISIFALGLYNGTKPLSTTSV 386
DB 303 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHONGIISILFALGLYNGTKPLSSTTA 362
QY 387 ESHEETDGYASWTFVFAARAYVEMMQCAKEPVLVRVLNDRVPLHGCANDKLGCRKR 446
DB 363 ENITQDGFSSAWTFVFAARAYVEMMQCQSEQEPVLVRVLNDRVPLHGPCVDALGCRCTR 422
QY 447 DDFVEGLSFARSGGNWAECEFA 467
DB 423 DSFVKGLSFARSGGDWGECEFA 443

RESULT 14
US-09-155-855-2
; Sequence 2, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANEKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0


```

; SEQ ID NO 2
;
; LENGTH: 443
;
; TYPE: PRT
;
; ORGANISM: Aspergillus niger
;
; FEATURE:
;
;   OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown
; US-09-155-855-2

```

Query Match	73.6%	Score 1813;	DB 4;	Length 443;
Best Local Similarity	76.0%;	Pred. NO. 1.6e-181;		
Matches 333; Conservative	46;	Mismatches 59;	Indels	0; Gaps 0;

QY	30	SCDVTGGVQCPPTSHLWGTYSPYFSLADESAISPDVPDCRVTFEQVLSRHGARYPTS	89
Db	6	TCDTVDQGYQCFSEPSHLWGQYAPFFSLANKSAISPDVPAGCHVTFAQVLSRHGARYPTD	65
QY	90	SASKAYSALIEAIOKNATAFKGYAFUKTYNYTLGADDLTPFCENOMWNSGKTFYRRYKA	149
Db	66	SGKKYSALIEIQONATTFGKYAFUKTYNSUGADDLTPFGEQLVNSGVKFQRYES	125
QY	150	LARKIVPTRASGSDRVIASAEKEIEGFQSAKLADPGSOPHOASPVINVIIEPGSGYNT	209
Db	126	LTRNIVPTRSSGSRVSIASGNKIEGFQSTKLKDPRAQPGQSSPKDIVVISEASTSNT	185
QY	210	LDHGCTCFADSEJLGDDEANFTALFAPATRAKLEADLPGLVTLDEDDVYVYLMDCMPFTV	269
Db	186	LDPGTCTVFDESELADDEIANFTATFVPSIQRLENLDSGLVTLDETVYVYLMDCSPDTI	245
QY	270	ARTSDATELSPFCALFTHDEWIOYDYLOSLGKYGYGACNPLGPAGQVGFANELLARLTH	329
Db	246	SNSTVDTKLSPCDLFTHEWINDYILOSLUNKYIGHGAGNPLGPTQGVGYANELIARLTH	305
QY	330	SPVQDHTCTNHTLDSNPATPLNATLYADFSDHNTMISITFFALGLYNGTKPLSTTSVESI	389
Db	306	SPVHDDTSSNHTLDSNPATPLNSTLYADFSDHNGIISILFALGLYNGTKPLSSSTAENI	365
QY	390	EBTDGYSASWTVPEAAAYVEMMQCAKEPLVRVLVNDRVVPLHGCACVDKIGCRCKRDF	449
Db	366	QTDTGFSASWTVPEASRMVYVEMMQCSEQEPVLRVLVNDRVVPLHGPCVPDALGCRTRDSF	425
QY	450	VEGLSFARSGGNWAECEFA	467
Db	426	VKGLSFARSGGDWGECEFA	443

RESULT 15

```

US-08-819-825-2
; Sequence 2, Application US/08819825
; Patent No. 5866118
;
; GENERAL INFORMATION:
;
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
;
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
;
; NUMBER OF SEQUENCES: 5
;
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 5400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435

```

```

, ATTORNEY/AGENT INFORMATION:
,
, NAME: Lambiris, Elias J.
,
, REGISTRATION NUMBER: 33,728
,
, REFERENCE/DOCKET NUMBER: 4758, 200-US
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 212 867 0123
,
, TELEFAX: 212 867 0298
,
, INFORMATION FOR SEQ ID NO: 2:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 475 amino acids
,
, TYPE: amino acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: protein
,
, FRAGMENT TYPE: internal
,
US-08-819-825-2

```

Query Match 54.28; Score 1334; DB 2; Length 475;
Best Local Similarity 56.18; Pred. No. 3e-131;
Matches 273; Conservative 62; Mismatches 114; Indels 38; Gaps 11;

QY	1	MGVF-VWLLSTATLFGSTCTALGP--RGNSHSCDVTGQYQCFPEIHLWGTSPYFSL	57
Db	6	LGSLFVLLQLQFSAII--TAGPAIPEFWRKHPNV-----IARHGWQSPFSL	52
QY	58	ADESAISPDVDDCRVTFVQVLSRHGARYPTSSAKAYASALIEAQKATAFKGKYAFLK	117
Db	53	AEVSEISPAVPGKRVFQVQVLSRHGARYPTAHKSEVAELLQRTQDTATFEKGDFALR	112
QY	118	TYNYTLGADDTLPTEGENQWNSGIKRYRYKALARKIYVPIRASGSDRVLTASAEKFTGG	177
Db	113	DYAYHLGADNLTRFCEBOMMESGRQFVHYRQAREIVPFVRAAGSARVLTAEFFNRGF	172
QY	178	QSAKLADPGSQPHQASPVINVIIPESGYNTLIDHGCTAFDESELGDVE-ANFTALFA	236
Db	173	QDAKDRDPRSNKQDAEVIINVISETGSNTLDGLTCAAA--EAPDPTQPAEFTQVFG	230
QY	237	PAIRARLEADLPGVTLTDEVDVYIMDCPFDTVARTSDAT---ELSPFCALTTHDEWQ	292
Db	231	PRVLKKTIKHMPGVNLTLEDVPLFMDLCPFDVVG--SDPVLPQLSPFCHLTADDWA	288
QY	293	VDYLSGLKYYGYGACGNPLPGQVGFANELLARLT-HSPVDHITSNTNHTLDSNPATFPL	351
Db	289	VDYVYTLKYYISHGGSAGFSGVGFVNELIARMTGLPKVDHTVYNTLDDNPETFL	348
QY	352	NATLYAFSDHNTMISIFFALGYNGTKPLSTVSE--SIEETQYGSASMTVPFAARV	409
Db	349	DAVLYAFSDHNTMTGIFSAMLYNGTKPLSTSKIPTGAAADYAASTVTFARAV	408
QY	410	EMMCK-----QAEKEPLRVLVNDRVVPLHGCAVDKLGRCCKRDDVEGLSFARG	459
Db	409	ELLRCETETSSEEEEDEEPFVRVLVNDRVVPLHGCRVDRWGRRCRRDEWIKGLTFARQ	468
QY	460	GNWABCF	466
Db	469	GHWDRCF	475

Search completed: October 26, 2001, 16:40:47
Job time: 4958 sec

QY	1141	ttgtctactactcttgtagaatctattgaagaacaaagcgggttaactctgcttcttgacct	1200
Db	1152	TTGTCTACTACTCTGTTGAATCTATTGAAGAACTGACGGTTACTCTGCTTCITGGACT	1211
QY	1201	gtccaccatcgctgtagagcttacattgaaatgatgcgaatgatgaagtcgaaaaggaaacca	1260
Db	1212	GTTCCATTTCGGTGCATAGAGCTTACGTTGAAATGATGCAATGTCGAAGCTGAAAAGGAACCA	1271
QY	1261	tttgttagagtgttggttaaagcacagagtgtgtccattgccagcgggttggtgttgacaag	1320
Db	1272	TGTGTTTAGATTGTTTAACGACACAGATGTTCCATTGCACGGTTGTGCTTGACAAG	1331
QY	1321	tgggttagatgaagagagacgacttcgttgaaaggtttgtctcttcgcgtagatctggtggt	1380
Db	1332	TTGGGTAGATGTAAGAGAGACGACTTCGTTGAAGGTTTGTCTTCGCCGTAGAICTGGTGGT	1391
QY	1381	aactgggaagaatgtttcgcttaa	1404
Db	1392	AAC TGGGCTGAATGTTTCGCTTAA	1415

RESULT	2	
AF295325		
LOCUS	AF295325	1350 bp mRNA SYN 18-OCT-2000
DEFINITION	Synthetic construct phytase mRNA, complete cds.	
ACCESSION	AF295325	
VERSION	AF295325.1	GI:10732782
KEYWORDS		
SOURCE	Synthetic construct.	
ORGANISM	Synthetic construct	
REFERENCE	artificial sequence.	
AUTHORS	1 (bases 1 to 1350)	
TITLE	Yang, L., Chen, Z., Bei, J., Liao, L. and Wang, X. Synthetic sequence of phytase gene for expression in Pichia pastoris	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 1350) Chen, Z.	
AUTHORS	Direct Submission	
TITLE	Submitted (11-AUG-2000) Animal Science Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, Guangdong 510640, China	
JOURNAL	3 (bases 1 to 1350)	
REFERENCE	Yang, L., Bei, J., Liao, L. and Wang, X.	
AUTHORS	Direct Submission	
TITLE	Submitted (11-AUG-2000) Biopharmaceutical Center, Zhongshan University, Guangzhou, Guangdong 510275, China	
JOURNAL	Location/Qualifiers	
FEATURES	1..1350	
source	/organism="synthetic construct" /db_xref="taxon:32630"	
CDS	/notes="Aspergillus niger sequence modified for expression in Pichia pastoris" 1..1350 /codon_start=1 /transl_table=11 /product="phytase" /protein_id="AAG22536.1" /db_xref="GI:10732783"	
	/translation="MLAVPASRNSQCDTVDOGYQCFTSHLWQYAPFFSLANESV ISPEVPAGCRVFVAQVLRSKHGKIPYDTSKGRKYSALIEEQNATTFDGKYAFLLTKYN YSLGADLLTPFGEQLVNSGIKFYORLESITRNIVPTFRSSGSRRVIASGKRIFEGQ PSIKRLDPAQPGCGSPKIDVVISSEASSNNITLDPGCTVTFEDELATVEANKPATAFV STLRKLENDLSGVTDITDPEVYLMDCMSFDITSTVDTKLSPFCDLPTHDEWTNYD YIQSLKKYYHGAGNPGLPGTCGVGANELIARLTSPVHDDTSSNHETLDSSPATFPFN STLYADEFSDNGIIISILFALGLYNGTKPLSTTVENITOTDGFSSANTVPPFAURLIVE MMQCAQEPLVRVLVNDVRVPELHCPCPDALGRCTNDSFVRGLSFARSGGDWAECA"	
BASE COUNT	319 a 354 c 282 g 395 t	
ORIGIN		

Query Match 55.8%; Score 783.6; DB 56; Length 1350;
Best Local Similarity 74.7%; Pred. No. 2e-198;
Matches 984; Conservative 0; Mismatches 334; Indels 0; Gaps 0;


```

/ gene="phyA"
prim_transcript 166..1871
/partial
/ gene="phyA"
/function="primary transcript of phyA gene"
sig_peptide join(210..253,356..365)
/ gene="phyA"
CDS join(210..253,356..1715)
/ gene="phyA"
/codon_start=1
/protein_id="CAA78904.1"
/db_xref="GI:2393"
/translation="MGVSAYLLPLYLLSGVTSLGAVPASRNQSCDVTDOGYCFSET
SHLWQVAPFIFSLANESVISPVPAGCRVTEAOVLSHGARYPTDSKGGKYSALIEI
QONATTGDKYAFUKTYNSIGADDLTPFGEQELVNSGIAFYQRYBSLTRNIVFIRS
SGSERVTSKSKFIEGQSKLKDPRAGQSSPKIDVISEASSNNTLDPTCTVF
EDSLADTAEANFTATVPISIRORLENDLSVLTIDTFTVYLMDSFDTISTVDT
KLSPFDLFTWINDYILQSLKKYKHGAGNPLGPTQGVYANELIARLTHSPVHD
DTSSNHTLDSPPATFPLNSTLYADFSDNGIIISILFALGLYNGTKPLSTTVENITQT
DGFSSAWTVFASRLYVEMMQCAEQEPLVRLVNDVPLHSCPVDALGRCTRDSFV
RGLSFARSGDWAECFA"
intron 254..355
/ gene="phyA"
/number=1
356..1878
/number=2
369..1712
/ gene="phyA"
1732..1760
/ gene="phyA"
BASE COUNT 463 a 603 c 456 g 478 t
ORIGIN

```

Query Match 43.5%; Score 610.2; DB 13; Length 2000;
 Best Local Similarity 64.9%; Pred. No. 4e-152;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

```

Qy 13 gtcgtgactgtccattgccacctgttgcgttccacatcccggtaccgcttgggtcct 72
Db 324 GTGGACTACTGATCGGTGACTATCTGTGCAGAGTCACTCCCGACTGGCGAGTCCCGCC 383
Qy 73 cgtgtaactctactctgtgacactgttgacggtgtggttaccagtgttccacagaatt 132
Db 384 TCGAGAAATCAATCCAGTTGCGATACGGTCGATCAGGGGTATCAATGCTTCTCCGAGACT 443
Qy 133 tctcaactgtgggtacatactctccattcttcttcttggctgacgaattctgtattct 192
Db 444 TCGCATCTTTGGGGTCAATACGACCCGTTCTCTCTGCGCAACGAATCGGTCACTCC 503
Qy 193 ccagacgttccaaagggttagagtaacttctgttcaagtttctgtagaacacggtgt 252
Db 504 CTTGAGTGTCCCGGATGACAGAGTCACTTCTGCTCAGGTCTCTCCCGTCAATGGAGCG 563
Qy 253 agatacccaacttctctgcgtctaaagggtactctgcttggattgaacttataaagg 312
Db 564 CGGTATCCGACCGACTCCAGGCGCAAGAAATATCTCGCTCAGTCTCTCCCGTCAATGGAG 623
Qy 313 aacgactgcttcaagggttagtagcgttttcttgaagacttacaactacactttgggt 372
Db 624 AACGCGACCACTTTGACGGAAATATGCTCTCTGCGCAACGAATCGGTCACTCCGCT 683
Qy 373 gctgacgactgactcattcgtgtagaacacaaatggttaactctgttataagttctac 432
Db 684 GCAGATGACCTGACTCCCTTCGGAGAACAGAGAGCTAGTCACTCCCGCATCAAGTTCTAC 743
Qy 433 agaagatacaagcttgggttagaagattgttccattcattagaacttctggttctgac 492
Db 744 CACCGGTACGAATCCGCTACAGGACACATCGTTCATTCCATTCCTCGATCCCTCGCTCCACG 803
Qy 493 agagttattgctctgctgaaagggttcattgaaggtttccaatctgtaagttgctgac 552

```

```

Db 804 CGCTGATCGCCTCCGCGCAAGAAATTCATCAGGGGCTTCAGAGCACCACCAAGCTGAAGGAT 863
Qy 553 ccaggtgctaacccacacacacagcttctccagttatitaaagttattattccaaaggtgct 612
Db 864 CTTGCTGCCCGAGCCCGGCAATCGTCCGCCAAGATCAGATCGTGTCTATTTCCGAGGCCAGC 923
Qy 613 ggttacaacaacacttggaccacggtttgtgtactgtcttccgaagaactctgaatgggt 672
Db 924 TCATCCAAACACACTCTCGACCCAGGACCTGCACCTGTCTTCGAAGACAGCAATTTGGCC 983
Qy 673 gacgaggtgaagcttaacttactgctgttttctgctccaccaattagactagattggaa 732
Db 984 GATACCGTCAAGCCCAATTTACACGCCACGCTGCTGCCCTCCATTCGTCAACGCTTGGAG 1043
Qy 733 gctcacttgcaggtgttaacttgaactgacgaagacgttgttaacttgaactgacacgtgt 792
Db 1044 AACGAGCTGTCGGGTGTGACTCTCAGACACAGAGTACCTACCTCATGACATGTGC 1103
Qy 793 cacttgcacactgttctagaaacttctgacgtactcaattgtctccattctgtgacttg 852
Db 1104 TCCTTCGACACCACTTCCACAGACCGTTCGACACCAAGCTGTCCCTCTCTGTGACCTG 1163
Qy 853 ttcaactcagcagaatggaatcaactgaactacttgaacttgaacttgggttaagtactacgt 912
Db 1164 TTCACCCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTGAAAAGATTACGGC 1223
Qy 913 tacggtgctggttaacccattgggtccagctcaaggtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1224 CATGGTGCAGGTAACCCGCTCGGCCGACCCAGGGGTCGGCTACGCTAACGAGCTCATC 1283
Qy 973 gctagattgactcaactctccagttcaagaccacacttctactaaacacacttggactct 1032
Db 1284 GCCGCTGTGACCACTCGGCTGTCCACGATGACACCAAGTTCCAAACACACTTTGGACTCG 1343
Qy 1033 aacccagctacttccattgaacgtacttgaacttgaacttgaacttctctcaacacacact 1092
Db 1344 AGCCCGGCTACCTTTCCGCTCAACTCTACTCTACAGGCACTTAAGCCGCTATCTACCAG 1403
Qy 1093 atggttctatttcttgcgttgggtttgttacaacggttactaaacgcttgcctactact 1152
Db 1404 ATCATCTCCATCTCTTTGCTTTAGGTCTGTACACAGGCACTTAAGCCGCTATCTACCAG 1463
Qy 1153 tctgttgaactctattgaagaactgaactgaactgctgtgtgtgtgtgtgtgtgtgtgtgt 1212
Db 1464 ACCGTTGGAATATATCACCCAGACAGATGGATCTGCTGTGCTTGGACGTTCCGTTGCT 1523
Qy 1213 gctagagcttactgttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1272
Db 1524 TCGCGTTTGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
Qy 1273 ttggttaacgacagaggttgcattgcaacggttgcattgcaacggttgcattgcaacggt 1332
Db 1584 TTGGTTAATGATCGGTTGTCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1643
Qy 1333 aagagagacgacttgcgttgaaggttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1392
Db 1644 ACCCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGTGATGCTGCTGCTGCTGCTGCTGCTG 1703
Qy 1393 ttttgcctta 1403
Db 1704 TGTGTTGCTTA 1714

```

RESULT 7

```

AR018076
LOCUS AR018076 2363 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 7 from patent US 5780292.
ACCESSION AR018076
VERSION AR018076.1 GI:3973679
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

```

Sat Oct 27 15:25:49 2001

REFERENCE 1 (bases 1 to 2363)
AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
Turkeli,T.K., Cantrell,M., Piddington,C.S., Rambosk,J.A.,
Turunen,M.K. and Fagerstrom,R.B.
TITLE Production of phytate degrading enzymes in trichoderma
JOURNAL Patent: US 5780292-A 7 14-JUL-1998;
FEATURES Location/Qualifiers
source 1..2363
BASE COUNT 559 a 732 c 510 g 562 t
ORIGIN /organism="unknown"

Query Match 43.5%; Score 610.2; DB 9; Length 2363;
Best Local Similarity 64.9%; Pred. No. 4e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgactgtccatgcccacctgttgcgttccacatccggtaccgcttgggtcct 72
Db 518 GTGGGACTACTGATCGCTGACAAATGTGCGAGATCACCCTCGGACTGGCAGTCCCGGCC 577

Qy 73 cgtgttaactctcaactctgtgacactgttgacggtgttaccatgtttcccgagaatt 132
Db 578 TCGAGAAATCAATCACCATTGCGATACGGTTCGATCAAGGTATCAATGCTTCTCGGAGACT 637

Qy 133 tctcaactgtgggtgacatactctccattctcttcttcttggctgacgaatctgctattct 192
Db 638 TCGCATCTTTGGGGTCAATACACCGCGGTCTTCTCTGCGCAACGAATCGCCATCTCC 697

Qy 193 ccagacgttcccaagggtgtgaggttactcttcttccagtttctgttagacacggtgtct 252
Db 698 CTTGATGTCCCGCGGTTCGAGATCACTTTTCGCTCAGGTCTCTCCCGTTCATGGAGCG 757

Qy 253 agatacccaactctctgtcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 312
Db 758 CGGTATTCGACCGAGTCCAAAGGCGCAAGAAATACCTCGCTCTCATTTGAGGAGATCCAGCAG 817

Qy 313 aacgtactgttccaaagggtgaagtaacgtcttcttgaagacttacaacttacttgggt 372
Db 818 AACGTGACCACTTTGATGGAAATATGCTTCTCTGAGACATACAACTACAGCTTGGGT 877

Qy 373 gctgacgactgtactccatctcgtgtgaacacaaatggttaactcgtggttattgaattctac 432
Db 878 GCAGATGACTGACTACTCCCTTCGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937

Qy 433 agaagatacgaagcttggctagaagattgttccattcattagagcttctggttctgac 492
Db 938 CAGCGATACGAATCGCTCACAGGAACATCATTCGTTTCATCCGATCTCTGGCTCCAGC 997

Qy 493 agagtattgtcttctgtgaaagttcattgaaggtttcccaatctcgttaagtgtggtgac 552
Db 998 CGCGTATCGCTCCGCGGAGAAATTCATTGAGGGCTTCAGAGCACCAGCTGAAGGAT 1057

Qy 553 ccagggtgttaaccac 612
Db 1058 CTTGTCGCCAGCGCGGCAATCTGTCGCCCAAGATGACGAGCTGCTTAATTCGAGGCGCAG 1117

Qy 613 ggttaacaaacacttggacacggttggtaactgttcttgaagaaactctgaattgggt 672
Db 1118 TCATCCAAACAACTCTGACCCAGGACCTGCACGTCTTTGAGACACAGCAATTGGCC 1177

Qy 673 gacgagttaagtaactcaactcgtgttcttgcgtcccaacttagagctagatggaa 732
Db 1178 GATACCGTCGAAGCAATTTACCGCGCAGCTTCGCGCCCTCCATTCGTCAAGCTCTGGAG 1237

Qy 733 gctcaactgtcgaaggtgttaactgtactgacgaagcgttggtaactgtgacatgtgt 792
Db 1238 AACGACCTGTCTGGGTAATCTCAACAGACACAGAGTACACCTACCTCATGNCATGTGC 1297

Qy 793 ccattgacactgtgtgagaactcttgacgctactcaattgttctccattctgtgactgt 852
Db 1298 TCCTTCGACACCATCTCCACACGACCGTTCGACACCAAGCTGTCCCGCTCTCTGTGACCTG 1357

Qy 853 ttaactcacgacgaatgatttcaatacagactacttgcgaatttgggtaagtactacggt 912
Db 1358 TTACCCCATGACGATGATGATCCACTACGACTACCTCCAGTCCCTGAAAAAATACTACGCG 1417

Qy 913 tacggtgtggttaac 972
Db 1418 CATGGCGCAGGTAAACCGCTCGCGGACGAGCGGCTGGGTACGCTAACGAGGTCACTC 1477

Qy 973 gctagattgactcaactctccagttcaagacacacacacacacacacacacacacacacacacac 1032
Db 1478 GCCCGTCTACCCACTCGCTGTCACGATGACACAGCTCCACACACACACCTTGGACTCG 1537

Qy 1033 aacccagctacttccacattgaacgctacttctgacgctgacttctctcaacacacacacacac 1092
Db 1538 AACCGAGTACCTTCGCGCTCAACTCTACTCTACGCGGACTTTTCCACGATAACGCG 1597

Qy 1093 atggttctatttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1152
Db 1598 ATCATCTATCTCTCTTCTTGGTCTGTACACGGCCTAAAGCGCTGTCTTACACAG 1657

Qy 1153 tctgtgaactctattgaagaaactgacggttactctgttcttcttcttcttcttcttcttcttct 1212
Db 1658 ACCGTGAGAAATATACCCAGACAGATGGTCTCTCTCTGCTTGGACGCTTCCGTTTGT 1717

Qy 1213 gctagagcttaccgttgaatgatgcaatgtgaagctgaaagaaacacacacacacacacacacac 1272
Db 1718 TCGCGTCTGTACGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777

Qy 1273 ttggttaacgacagagattgttccattgacggttgggttgggttgggttgggttgggttgggt 1332
Db 1778 TTGGTTAATGATCGGCTGTGCGGCTGATGCGGTGTCCTGCTTCCAAATGATGCTTTTGGGAGATGT 1837

Qy 1333 aagagacgactctgttgaaggttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1392
Db 1838 ACCCGGATAGCTTTGAGGGGTTGAGCTTTGCTAGATCTGGGGTGAATGGCGGAG 1897

Qy 1393 tgttctgctta 1403
Db 1898 TGTCTGCTTA 1908

RESULT 8
AR051916 2363 bp DNA PAT 29-SEP-1999
LOCUS Sequence 7 from patent US 5830733.
DEFINITION AR051916
ACCESSION AR051916
VERSION AR051916.1 GI:5975280
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2363)
AUTHORS Nevalainen,H.K.M., Paloheimo,M.T., Miettinen-Oinonen,A.S.K.,
Turkeli,T.K., Cantrell,M., Piddington,C.S., Rambosk,J.A.,
Turunen,M.K., Fagerstrom,R.B. and Houston,C.S.
TITLE Nucleic acid molecules encoding phytase and pH2.5 acid phosphatase
JOURNAL Patent: US 5830733-A 7 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..2363
BASE COUNT 559 a 732 c 510 g 562 t
ORIGIN /organism="unknown"

Query Match 43.5%; Score 610.2; DB 9; Length 2363;
Best Local Similarity 64.9%; Pred. No. 4e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgactgtccatgcccacctgttgcgttccacatccggtaccgcttgggtcct 72
Db 518 GTGGGACTACTGATCGCTGACAAATGTGCGAGATCACCCTCGGACTGGCAGTCCCGGCC 577

Qy 73 cgtgttaactctcaactctgtgacactgttgacggtgttaccatgtttcccgagaatt 132
Db 578 TCGAGAAATCAATCACCATTGCGATACGGTTCGATCAAGGTATCAATGCTTCTCGGAGACT 637

Qy 133 tctcaactgtgggtgacatactctccattctcttcttcttggctgacgaatctgctattct 192
Db 638 TCGCATCTTTGGGGTCAATACACCGCGGTCTTCTCTGCGCAACGAATCGCCATCTCC 697

Qy 193 ccagacgttcccaagggtgtgaggttactcttcttccagtttctgttagacacggtgtct 252
Db 698 CTTGATGTCCCGCGGTTCGAGATCACTTTTCGCTCAGGTCTCTCCCGTTCATGGAGCG 757

Qy 253 agatacccaactctctgtcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 312
Db 758 CGGTATTCGACCGAGTCCAAAGGCGCAAGAAATACCTCGCTCTCATTTGAGGAGATCCAGCAG 817

Qy 313 aacgtactgttccaaagggtgaagtaacgtcttcttgaagacttacaacttacttgggt 372
Db 818 AACGTGACCACTTTGATGGAAATATGCTTCTCTGAGACATACAACTACAGCTTGGGT 877

Qy 373 gctgacgactgtactccatctcgtgtgaacacaaatggttaactcgtggttattgaattctac 432
Db 878 GCAGATGACTGACTACTCCCTTCGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937

Qy 433 agaagatacgaagcttggctagaagattgttccattcattagagcttctggttctgac 492
Db 938 CAGCGATACGAATCGCTCACAGGAACATCATTCGTTTCATCCGATCTCTGGCTCCAGC 997

Qy 493 agagtattgtcttctgtgaaagttcattgaaggtttcccaatctcgttaagtgtggtgac 552
Db 998 CGCGTATCGCTCCGCGGAGAAATTCATTGAGGGCTTCAGAGCACCAGCTGAAGGAT 1057

Qy 553 ccagggtgttaaccac 612
Db 1058 CTTGTCGCCAGCGCGGCAATCTGTCGCCCAAGATGACGAGCTGCTTAATTCGAGGCGCAG 1117

Qy 613 ggttaacaaacacttggacacggttggtaactgttcttgaagaaactctgaattgggt 672
Db 1118 TCATCCAAACAACTCTGACCCAGGACCTGCACGTCTTTGAGACACAGCAATTGGCC 1177

Qy 673 gacgagttaagtaactcaactcgtgttcttgcgtcccaacttagagctagatggaa 732
Db 1178 GATACCGTCGAAGCAATTTACCGCGCAGCTTCGCGCCCTCCATTCGTCAAGCTCTGGAG 1237

Qy 733 gctcaactgtcgaaggtgttaactgtactgacgaagcgttggtaactgtgacatgtgt 792
Db 1238 AACGACCTGTCTGGGTAATCTCAACAGACACAGAGTACACCTACCTCATGNCATGTGC 1297

Qy 793 ccattgacactgtgtgagaactcttgacgctactcaattgttctccattctgtgactgt 852
Db 1298 TCCTTCGACACCATCTCCACACGACCGTTCGACACCAAGCTGTCCCGCTCTCTGTGACCTG 1357

Db 578 TCAGAAATCAATCCACTTCGATACGGTCGATCAAGGTATCAATGCTTCCGAGACT 637
Qy 133 tctcaactgtgggtacataactctccattctctctctgttggtgacgaatctgtctattct 192
Db 638 TCGATCTTTGGGGTCAATACGCGCGTCTCTCTCTGCGAAAGAAATCGCCATCTCC 697
Qy 193 ccagacgttccaaaggtgttagaacttctcgttcaagtttctctagacacgggtct 252
Db 698 CCTGATGTGCGCGCGGTTCGAGAGTCACTTCGCTCAGGTCTCTCCGTCATGGAGCG 757
Qy 253 agataccaaactctctctgtctctaaagcgctactctctgtttgattgaagctattcaaaag 312
Db 758 CGGTATCCGACCGAGTCCAAAGGCAAGAAATACTCCGGTCTCATTTGAGGAGATCCAGCAG 817
Qy 313 aacgctactgctttcaaggglaagtaagcttcttgaagacttaacaaactaactttgggt 372
Db 818 AACGTGACCACTTTGATGGAATAATATGCCCTCTGTAAGACATACAACCTACAGCTTGGT 877
Qy 373 gctgacgactgactccattcgggtgaacaaatggttaactctgtgtattaaattctac 432
Db 878 GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC 937
Qy 433 agaagatacaaggttgggttagaagattgttccattccattagagcttctgtttctgac 492
Db 938 CAGCGATACGAATCGCTCACAGGAACATCATTCCTTCATCCCATCTCTGGCTCCAGC 997
Qy 493 agagtattgtctgtgaaagtcattgaagtttccaaatctgtaagttggtgac 552
Db 998 CGGTGATGCGCTCGGGGAGAAATCATTTGAGGGCTTCAGAGCACCAAGCTCAAGGAT 1057
Qy 553 ccaggtgtaaccacacacacagcttctccagttatttaacgttatttccagaaggtgct 612
Db 1058 CCTCGTCCGACCGCGGCAATCTGTCGCCCAAGATCAGCTGTCATTTCCGAGGCCAGC 1117
Qy 613 gtttaacaaacacttggaccacaggttctgtgactgcttccaaagaaatctgaattgggt 672
Db 1118 TCATCCAAACACTCTCGACCCAGGACCTGCACTGTCTTTGAAGACAGCAATTTGGCC 1177
Qy 673 gacgactgtgaagtaacttcaacttcaactgtgttttgcctccacaaatagagctagatggaa 732
Db 1178 GATACCGTCGAAGCAATTTACCGCCACGCTTCGCCCCCTCCCATTCGTCAACGCTGGAG 1237
Qy 733 gctcaactgcaaggtgtaacttgaactgacgaagcgttgttaactgtgagacatgtgt 792
Db 1238 AACGACCTGTCTGCGGTGACTCTCACAGACAGAGTGAACCTACCTCATGAGCATGTGC 1297
Qy 793 ccatcgaactgtgtgtagaacttctgaagctactcaaatgtctccattctgtgacttg 852
Db 1298 TCCTTCGACACCATCTCCACGACGCGTCGACACCAAGCTGTCCGCCCTCTCTGTGACCTG 1357
Qy 853 ttaactcagcagcaatggattcaactgactacttgaactcttcaatcttgggttaagtactacgt 912
Db 1358 TTCACCCATGACGAATGGATCCACTAGGACTACTCCAGTCCCTGAAAATACTACCGGC 1417
Qy 913 taagggtgtaaacacttgggtccagctcaaggtgtgttgggttcttaacgaattgatt 972
Db 1418 CATGGCGAGGTAAACCGCTTCGCGCCGACCCAGGGGTGCGGTACGCTAACGAGCTCATC 1477
Qy 973 gctagattgactcactctcagttcaagaccacacttctactaacaacaaacttggactct 1032
Db 1478 GCCGCTCACCCACTTCGCTGTCCACGATGACACCGCTTCAACACACACCTTGGACTCG 1537
Qy 1033 aaccagctacttccacttgaacgctacttgttagcgtgactctctcactcagacaact 1092
Db 1538 AACCCAGCTACCTTCGCGTCAACTCTACTCTACGCGGACTTTTCCACGATACGGC 1597
Qy 1093 atggtttctatttcttctgttgggttgaacagcgttaacgaacttctactact 1152
Db 1598 ATCATCTCATCTCTTCTTGGGTCTGTATACAGGCACTTAAGCGGTGTCTACCGC 1657
Qy 1153 tctgttgaactctattgaagaaactgacttactctgtcttctgactgttccattcgtct 1212

Db 1658 ACCGTGAGAAATATCAACCACAGACAGATGGCTTCTCTGCTTCGACGGTTCGGTTGCT 1717
Qy 1213 gctagagcttacgttgaatgatgcaatgtgaagctgaaagaaacacattggttagagtt 1272
Db 1718 TCGGTCTGTACGTGAGATGATGCACTGCAGCGCCAGCAGGCGCTGGTCCGTGTC 1777
Qy 1273 ttggttaacagacagagttgttccattgacggttgggtggtgacaaagttgggttagatgt 1332
Db 1778 TTGTTAATGATCGCTTGTCCGCTGCATGGGTGCCAATTGATGCTTTGGGAGACT 1837
Qy 1333 aagagagacgactcttgaagtttcttctctgctagatctggttggttaactggaagaa 1392
Db 1838 ACCGGGATAGCTTTGTGAGGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTTGGCGGAG 1897
Qy 1393 tgtttcgctta 1403
Db 1898 TGTCTGCTTA 1908

RESULT 9
AR053934
LOCUS AR053934 2379 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5834286.
ACCESSION AR053934
VERSION AR053934.1 GI:5978796
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2379)
AUTHORS Nevalainen,H.K.M., Palohelmo,M.T., Fagerstrom,R.B.,
Miettinen-Oinonen,A.S.K., Turunen,M.K., Rambosk,J.A.,
Piddington,C.S., Houston,C.S. and Cantrell,M.A.
TITLE Recombinant cells that express phytate degrading enzymes in desired
ratios
JOURNALS Patent: US 5834286-A 1 10-NOV-1998;
FEATURES Location/Qualifiers
source 561 a 735 c 518 g 565 t
BASE COUNT 561 a 735 c 518 g 565 t

Query Match 43.5%; Score 610.2; DB 9; Length 2379;
Best Local Similarity 64.9%; Pred. No. 4e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgctactgtccattgcccacttgttcggttccacatccggtacogccttgggtct 72
Db 534 GNGGGACTACTGATCGCTGACAAATCTGCGATACGGTTCGATCAAGGGTATCAATGCTTCCGAGACT 593
Qy 73 cgtggttaactctactctgtgacactgttgacggtggttaccatgtttcccaagaatt 132
Db 594 TCGAGAAATCAATCCACTTGCATACGGTTCGATCAAGGGTATCAATGCTTCCGAGACT 653
Qy 133 tctcaactgtgggttacataactctccattctcttcttggctgacgaatctgtattct 192
Db 654 TCGCATCTTTGGGTCAATACGCGCGCTTCTCTCTCTGCGAAACGAATCGCCATCTCC 713
Qy 193 ccagacgttccaaaggtgttagaacttctcgttcaagtttctctagacacgggtct 252
Db 714 CCTGATGTGCGCGCGGTTCGAGAGTCACTTTCGCTCAGGTCTCTCCGTCATGGAGGG 773
Qy 253 agataccaaactctctctgtctctaaagcgctactctgttcttgaagctattcaaaag 312
Db 774 CGGTATCCGACCGAGTCCAAAGGCAAGAAATACTCCGCTCTCATTTGAGGAGATCCAGCAG 833
Qy 313 aacgctactgtttcaagggtaagtaagcttcttcttgaagacttaacaaacttgggt 372
Db 834 AACGTGACCACTTTGATGGAATAATATGCCCTTCTTGAAGACATACAACCTACAGCTTGGT 893
Qy 373 gctgacgacttgaactcctcgttgaacaaacaaatggttaactcgttattgaagttctac 432

Db 894 GCAGATGACCTGACTCCCTTCGGAGAGAGAGAGAGAGTACTCAACCCGCGCATCAAGTTCTTAC 953
 Qy 433 agaagatacaaggcttggcttagaaaaattgttccattcattagagcttctgtgttgac 492
 Db 954 CAGCGATACGAATCGCTCACAGAACATCATTCGTTTCACTCCGATCCTCTGGCTCCAGC 1013
 Qy 493 agagtattgtctgtgtaaaagttaattgaagtttccaaatctgtaagtgtgctgac 552
 Db 1014 CCGGTGATCGCTCCGGGAGAAAATTCATTGAGGGCTTCAGAGACCAACAGCTCAAGGAT 1073
 Qy 553 ccaggtgtaaccacacacagctctccagttattaaacgtttatttccagaaaggtgct 612
 Db 1074 CTTCTGTCACCGCGGCAATCTGCGCCCAAGATCGACGTGCTCATTCGAGGCCAGC 1133
 Qy 613 ggttaacaaacacactttggaccaggttggtagtactgttctcgagaatctgtaattgggt 672
 Db 1134 TCATCAACACACTCTCGACCCAGGCACTGCACTGTCTTTGAGACAGCGCAATTGGCC 1193
 Qy 673 gacgagttgaagtaacttcaactctgctgttttcgctccacaaattagagctagatgaa 732
 Db 1194 GATACCGTCGAAGCAATTCACCGCCACAGTTCGCCCTCCATTCGTCACAGCTCGAG 1253
 Qy 733 gctcacttgccaggtgttaacttgactgacgaagacgtgttgaacttgatgacatggt 792
 Db 1254 AACGACCTGTCTGGCGTGACTCTCACAGACACAGAAAGTACCTCATGGACATGTC 1313
 Qy 793 ccattgacactgttgcagaaactctgacgtactcaatgttccattctgtgacttg 852
 Db 1314 TCGTTTCGACACCATCTCCACAGCACCGCTCGACACACCAAGCTGTCCCTTCTGTGACCTG 1373
 Qy 853 ttacactcagcgaatgatttaatacactacttgcgaatcttgggttaagtactacgtt 912
 Db 1374 TTCACCATGACGATGATGATCACTACGACTACCTCCAGTCCCTGAAAAAATACTACGGC 1433
 Qy 913 tacggtgtggttaacccattgggtccagtcacgaaggtgtggtttggttaacgaattgatt 972
 Db 1434 CATGGCGCAGGTAAACCGCTCGCGCCGACCCAGGCGCTGGGTACGCTAACGAGCTCATC 1493
 Qy 973 gctagattgaactcactctccagttcaagacacacttctactaacacactttgactct 1032
 Db 1494 GCCGCTTCACCCACTCGCCTGTCCACGATGACACAGCTCCACACACACTTGGACTCG 1553
 Qy 1033 aaccagctacttccactgaacgtacttctgtacgtgacttctctcaagacacact 1092
 Db 1554 AACCAGCTACCTTCGCGCTCAACTCTACTCTTACGCGGACTTTTCCACGATAACGCG 1613
 Qy 1093 atggtttctatttctcgtttggtttgttcaacaggttactaaagccattgtctactact 1152
 Db 1614 ATCATCTATCCTCTTCTGCTTGGCTGTGTACAACGCGCCTAAGCCGCTGTCTACCCAG 1673
 Qy 1153 tctgtgaactctattgaagaaactgacgttactctgtcttctgtgactgttccattcgt 1212
 Db 1674 ACCGTGGAGAAATATCACCAGACAGATGGGTCTCTGCTCTGTGACGCTTCCGTTGCT 1733
 Qy 1213 gctagagcttactgtgaaatgatgcaatgtgaagctgaaagaaacattggttagagtt 1272
 Db 1734 TCGCTGTGTACGTGAGATGATGTCAGTGCAGCGCGGAGCAGGAGCGCTGCTGCTGTC 1793
 Qy 1273 ttggttaacgacagagattgttccattgcaaggttgggttggtgacaaagtgggttagatg 1332
 Db 1794 TTGGTTAATGATCGCTGTCCCGCTGTCATGGGTGTCCAATGATGCTTTGGGGAGATGT 1853
 Qy 1333 aagagacgactctgtgaaagtttcttcttcgctagatctggttgtaactgggaagaa 1392
 Db 1854 ACCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGCGGAG 1913
 Qy 1393 tgtttgcctta 1403
 Db 1914 TGTCTGCTTA 1924

RESULT 10
 ASPHYTAS

LOCUS ASPHYTAS 2379 bp DNA PLN 23-MAR-1994
 DEFINITION Aspergillus niger var awamori phytase gene, complete cds.
 ACCESSION L02421
 VERSION L02421.1 GI:166518
 KEYWORDS phytase.
 SOURCE Aspergillus niger (strain ALK0243, sub-species awamori) DNA.
 ORGANISM Aspergillus niger
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 REFERENCE 1. (sites)
 AUTHORS Piddington,C.S., Houston,C.S., Paloheimo,M., Cantrell,M.,
 Miettinen-Oinonen,A., Nevalainen,H. and Rambosek,J.
 TITLE The cloning and sequencing of the genes encoding phytase (phy) and
 pH 2.5-optimium acid phosphatase (aph) from Aspergillus niger var.
 awamori
 JOURNAL Gene 133 (1), 55-62 (1993)
 MEDLINE 94040796
 REFERENCE 2 (bases 1 to 2379)
 AUTHORS Carter,J.R., Franden,M.A., Aebersold,R.H. and McHenry,C.S.
 TITLE Molecular cloning, sequencing and overexpression of the gene
 encoding the psi subunit of E. coli DNA polymerase III holoenzyme
 Unpublished (1992)
 JOURNAL Location/Qualifiers
 FEATURES
 Source
 1. .2379
 /organism="Aspergillus niger"
 /db_xref="taxon:5061"
 420..463
 /citation={2}
 /evidence=experimental
 join(420..463,566..1925)
 /EC_number="3.1.3.8"
 /citation={2}
 /codon_start=1
 /function="conversion of phytate to inositol and inorganic
 phosphorus"
 /evidence=experimental
 /product="phytase"
 /protein_id="AA16898.1"
 /db_xref="GI:166519"
 /translation="MGSAVLPLLYLLAGVTSGLAVPASRNQSTCDTVDQVQCFSET
 SHLQGYAPFSLANESALSDPVACGRVTFQVLSRHGARYPTESKGGKYSALLEEI
 QNVTFDGYAFLKTYNSLGDLLTPFGEOLVNSGKIFQRYESLIRNIPFIRS
 SSSRVIASGEKFIEGFOSYKLDKDPRAOPGSSPKLDVVISSESSNNTLDPTCTVF
 ESELDVTVEANFTATFAPSISQRLENDLSGVTLTDTETVYLMDCMSPDTISTSVDT
 KLSPPCDLTHDEWIHYDLSLKYIHGAGNPLGPTGCGVYANCLARLHLSVPHD
 DTSNFDLNSFATPLNSTLADFSDHNGIISILFAUGLYNGTLPSTTTVENTITQT
 DGFSSAWTVPFASRLYVEMQCAEQEPLRVLVNDRVPLVPLHGCPIDALGRCRDSFY
 RGLSFARSGGDWAECSA"
 566..1925
 /note="putative"
 /citation={2}

BASE COUNT 561 a 735 c 518 g 565 t
 ORIGIN
 Query Match 43.5%; Score 610.2; DB 13; Length 2379;
 Best Local Similarity 64.9%; Pred. No. 4e-152;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;
 Qy 13 gtcgtgctactgtccattgcccacctgtgtcggttccacatccggtaccgcttgggtccct 72
 Db 534 GTGGGACTACTGATCGCTGACAACTCTGTGACAGATCACTCCGAGCTGGCAGTCCCGGCC 593
 Qy 73 cgtggttaactctactcttctgacactgttgacggttggttaccgaatttccacgaatt 132
 Db 594 TCGAGAAATCAATCACTTGGATACGTCGATCAAGGGTATCAATGCTTCTCGAGACT 653
 Qy 133 tctcaactgtgggttacatactctccattctctctgtgctgacgaattctgtattct 192
 Db 654 TCGCATCTTTGGGGTCAATACGCGCGGTCTTCTCTCTCGAAACGAATCGGCATCTCC 713
 Qy 193 ccagacgttccaaaggggtgttagagttacttctcgaagtcttctagacacaggtgct 252

Db	857	TCGAGAAATCAATCAGTGGATACGGTCGATCAGGGGTATCAATGCTTCCTCCGAGACT	916
Qy	133	tctcaactgtgggttacatactctccattctctctcttcttggctgacgaatctgctattctt	192
Db	917	TGCAATCTTTGGGGTCAATACGACCGTCTCTCTCTCTGGCAAGGAATCGTCACTCC	976
Qy	193	ccagacgtttcccaagggttagagttactcttcgttccaaagtttgcctagacacaggtgct	252
Db	977	CCTGAGTGCCTCCCGGATGCGAGTCACTTTCGCTCAGGTCCTCTCCCGTCATGGAGCG	1036
Qy	253	agatacccaactctctctgctctgaagcgctactctgcttggattgagactattcaaaag	312
Db	1037	CGGTATCCGACCGACCCCAAGGGCARGAANAATACCTCCGCTCATTTGAGGAGATCCAGCAG	1096
Qy	313	aacgctactgctttccaagggttaagtcacgctcttcttgaagacttaacaactacactttgggt	372
Db	1097	ARGCGACCACTTTGACGGAAATATGCTTCCTGAAGACATACAACCTACAGCTTGGGT	1156
Qy	373	gctgacgacttgaactccattcgttgtaacacaaatgggttaactctggtattaagtctac	432
Db	1157	GCAGATGAGCTGACTCCCTTCGGAGAACAGAGGAGCTAGTCAACTCCGGCATCAAGTTCAC	1216
Qy	433	agaagatacaaggcttggctgataaagattgtccattcattagagcttctggttctgac	492
Db	1217	CAGCGGTACGAATCGCTCACAGGACATCGTTCCATTCATCCGATCCTCTGGCTCCAGC	1276
Qy	493	agagttattgcttctgctgtaaaagttcattgaaggtttccaatctgcttaagtgtgctgac	552
Db	1277	CGCGTGATCGCCCTCGCGCAAGAAATTCATCAGGGGCTTCAGAGCACCAAGCTGAAGAT	1336
Qy	553	ccaggtgtctaaccacacacaaagttctcaagttattaaagcttattatccaaaggtgct	612
Db	1337	CCTCGTCCGACGCGCCCAATCGTCGCCCAAGATCGACGCTGCTCAITTTCCGAGGCCACG	1396
Qy	613	ggttacaacaacacttggaccacggtttgtgtaactgcttccaaagaaactctgaattgggt	672
Db	1397	TCATCCAAACAACTCTCGACCCAGGACCTGCACGTCTTCGAAGACAGCAATTTGGGC	1456
Qy	673	gaacggtgaagtaacttaacttcaactgtgttttgcgtcccaaatagagactagattgaa	732
Db	1457	GATACCTCGAAGCAATTTCAACCGCACGCTTCGTCCTCCCTCCATTCGTCAACGCTCGGAG	1516
Qy	733	gctcaactgccaaggtttaacttgaactgacgaagcgttgttaacttgaactgacatgtgt	792
Db	1517	AACGACCTGTCCGGTGTGACTCTCACAGACACAGAAGTGAACCTACCTCAAGGACATGTC	1576
Qy	793	ccattcgacactgtgtctagaactcttgacgctactcaattgtctctccattctgtgacttg	852
Db	1577	TCCTTCGACACCATCTCCACCGACCGTTCGACACCAAGCTGTCCCTCTCTGTGACCTG	1636
Qy	853	ttcactcaacgaagaatgattcaatacagactacttgcacttcttgggttaagtaactcaggt	912
Db	1637	TTCAACCATGCAATGGATCAACTAGACTACTCTCCAGTCTCCAGTCTTCAAGAAAGTATTACGGC	1696
Qy	913	tacggtactgtaaccacttgggtccagctcaaggttgttgggttgcgttaacgaattgatt	972
Db	1697	CATGGTCAGGATAACCCGCTCGGCCCGACCCAGCGGCTCGGCTACGCTACAGACTCAIC	1756
Qy	973	gctagattgactcactctccagttccaagacacacacttctactaaccacactttggactct	1032
Db	1757	GCCGCTGTGACCCACTCGCTGTCCAGGATGACACCAAGTTCCCAACCACTTTGGACTCG	1816
Qy	1033	aaccagactacttcccatgaacgctacttgaagctgaacttctctcactcagcaaacact	1092
Db	1817	AGCCCGGTACTCTTCCGCTCAACTCTACTCTCTCAGCGGACTTTTCGATGACAGCGC	1876
Qy	1093	atggttctatttctctgcttgggtttgtacaacggttactaagccattgtctactact	1152
Db	1877	ATCATCTCCATCTCTTGTCTTTAGTCTGTACAAACGGCAAGCGGCTATCTACCAAG	1936
Qy	1153	tctgtgaatcatttgaagaaactgaagcgttactctgtcttcttggactgttccattgct	1212
Query Match			
43.5%; Score 610.2; DB 9; Length 6756;			

Db	1937	ACCGTGAGAAATATACCCAGACAGATGATCTCTGCTCTGGACGGTTCGGTTGCT	1996
Qy	1213	gctagagcttacgttgaatgatgcaatgtgaagctgaaagaaacacattggttagagtt	1272
Db	1997	TCGCGTTTGTACCTCAGATGATGTCAGTGTTCAGCGGACGAGACGCCGTGTCGGTGC	2056
Qy	1273	ttggttaacgacagagttgttccattgcaaggttgggtgttgacaaagttgggttagatgt	1332
Db	2057	TTGGTTAATGATCGCTGTCTCCGCTGCATGGGTGTCGGTTGATCTTTGGGGAGATGT	2116
Qy	1333	aaagagacacttcgttgaaggtttgtcttctgctagatctggtgtaacttgggaagaa	1392
Db	2117	ACCGGATAGCTTTGTAGGGGTTGAGCTTTGCTAGATCTGGGGGTGATTGGCGGAG	2176
Qy	1393	tggttcgctta	1403
Db	2177	TGTTTGTCTTA	2187
RESULT 12			
LOCUS A19451 6756 bp DNA PAT 10-JUN-1994			
DEFINITION phytase gene locus of plasmids pAF 2-3, pAF 2-6, pAF 2-7.			
ACCESSION A19451			
VERSION A19451.1 GI:583193			
KEYWORDS synthetic construct.			
SOURCE synthetic construct.			
ORGANISM synthetic construct			
REFERENCE 1 (bases 1 to 6756)			
AUTHORS Veenstra,A.E., Luiten,R.G.M. and Seiten,G.C.M.			
TITLE Cloning and expression of microbial phytase			
JOURNAL Patent: EP 0420358-A 40 03-APR-1991;			
GIST-BROCADES N.V			
FEATURES			
Location/Qualifiers			
1..6756			
/organism="synthetic construct"			
/db_xref="taxon:32630"			
join(210..254,357..1715)			
/gene="phytase gene"			
join(210..254,357..1715)			
/gene="phytase gene"			
join(210..254,357..1712)			
/gene="phytase gene"			
210..254			
/gene="phytase gene"			
number=1			
join(210..254,357..1715)			
/gene="phytase gene"			
/codon_start=1			
/transl_table=11			
/protein_id="CAA01463.1"			
/db_xref="GI:583194"			
/translation="MGVSAVLLPLLLSGVTSGIAPVSPASRNSQSCDVGOCFSET			
SHLWQYAPFESLANESVISPVPAGCRVTFQVLSRHGARYPTDSKGYKYSALIEI			
QQAATTFDGYAFILKTYNSLGADLTTFGEQELVNSGKIFQYRISEASSNNTLDPGTIVF			
SGSRVIAISGKKFTEGFQSKLKDPRAPQGSPPKIDWVISEASSNNTLDPGTIVF			
EDSLADTVEANFTATVPISIRORLENDLSGVTITDTEVTYLMDCMSFDTISTVDT			
KLSPCDLFTHDWINYDYLQSLKKYHGHGAGNPLGPTGVGYANELLARLTHSPVHD			
DTSNHLIDSSPATFPLNSTLYADFSDHNGIISILFALGLYNGTKPLSTTVENITQT			
DGFSSAMTVFASRLYEMVQCEQEPLVRLVNDVRVPLHGPCVDALGRCTDSEFV			
RGLSFARSGGDWAECFA"			
255..356			
/number=1			
357..1715			
/gene="phytase gene"			
/number=2			
BASE COUNT 1615 a 1712 c 1712 g 1714 t 3 others			
ORIGIN			

Best Local Similarity 64.9%; Pred. No. 4.3e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

```
QY 13 gtcgtactacttccatgacacacttgcgttccacatccggtacccgcttgggtcct 72
Db 324 GTGGGACTACTGATCGCTGACTATCTGTGACAGAGTCACTCCGGACTGGCAGTCCCGCC 383
QY 73 cgtgtaactctcaactcttctgtgacactgttgacggtgttaccatggtttcccaagaatt 132
Db 384 TCGAAGAAATCAATCCAGTTCGATACGATACGATACGATACGATACGATACGATACGAT 443
QY 133 tctcactgtggtggtacacacttccatcttcttcttcttcttcttcttcttcttcttctt 192
Db 444 TCGCATCTTTGGGGTCAATACGACACCGTCTCTCTCTGGCAACGAATCGTATCTCC 503
QY 193 ccagcgtctccaaagggtgttagaacttcttcttcttcttcttcttcttcttcttcttctt 252
Db 504 CCGTAGGTGCCCGCGGATGACAGAGTCACTTTCGCTCAGGTCTCTCCCGCTCAATGGAGCG 563
QY 253 agatacccaacttcttctgcttaagcgttactcttcttcttcttcttcttcttcttcttctt 312
Db 564 CGGTATCCGACCGACTCCCAAGGCAAGAAATACTCCGCTCTCATGTAGGAGATCCAGCAG 623
QY 313 aacgtactgttctcaagggttaagtaagcttcttcttcttcttcttcttcttcttcttctt 372
Db 624 AACGGACACCTTTGACGGAAATATGCTTCTCTGAGACATACAACTACAGCTTGGGT 683
QY 373 gctgacgacttgaactccatctcgttgaacacacaaatgttaactctgttatttaagtctac 432
Db 684 GCAGATGACCTGACTCCCTTCGGAGACAGAGTACTGTAACCTCCGGCATCAAGTTCTAC 743
QY 433 agaagatacagaaggttgggttagaagattgttccatcttcttcttcttcttcttcttctgac 492
Db 744 CAGCGGTACGAATCGCTCACAGGAACATCGTTCATTCATCCGATCCTCTGGGTCCAGC 803
QY 493 aaggtattgtctctgctgaaggttcaattgaaggttccaaatcttcaagttgggtgac 552
Db 804 CGCTGATCGCTCCGGACAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGAT 863
QY 553 ccaggtgttaaccacacacacacacacacacacacacacacacacacacacacacacacacac 612
Db 864 CCGTGTGCCAGCGCGCAATCTGCGCCCAAGATCGACGTGTCTATTCGGAGGCCAGC 923
QY 613 gtttacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 672
Db 924 TCATCCCAACACACTCTCGACCCAGCACCTGACCTGTCTTCAAGACACAGCAATTTGGCC 983
QY 673 gacgacttgaagcttaacttcaactgtgttctgctccacacacacacacacacacacacacac 732
Db 984 GATACCGTCGAAGCAATTTACCGCCACGTTGTCGCCCTCCATTCGTCAACGTCCTGGAG 1043
QY 733 gctcaactgtccaggtgttaacttgaactgacgaagcgttgaacttgaacttgaacttgaact 792
Db 1044 AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTGAACCTACCTCATGGACATGTGC 1103
QY 793 ccattgacacactgtgttagaactctgactgaactgaactgaactgaactgaactgaactgaact 852
Db 1104 TCCTTCAGACACATCTCCACACACACCGTTCGACACCAAGCTGTCCCGCTTCTGTGACCTG 1163
QY 853 ttcactacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 912
Db 1164 TTCACCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTTGAAAGATTTACGGC 1223
QY 913 tacggtgtgttaaccatgtgtccagctcaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1224 CATGGTGCAGGTAAACCGCTTCGGCCGAGCCAGCGGCTCGGCTACGCTAACGAGCTCATC 1283
QY 973 gctagatgactactctccagttcaagcaccacacttctactaaccacacacacacacacactt 1032
Db 1284 GCCCGTGTGACCCACTCGCTGTCCAGATGACACCGATTCACACGATTCACACGATTTGAC 1343
QY 1033 aaccagctacttccattgaacgctacttctgtacgctgaacttcttctcaacacacact 1092
```

```
Db 1344 AGCCCGGCTACCTTTCGGCTCAACTCTACTCTCTACCGGACTTTTCGCATGACAAGGC 1403
QY 1093 atggtttctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1152
Db 1404 ATCATCTCCATCTCTTGTGTTAGTCTGTACACGGCACTAAGCGCTATCTACACAG 1463
QY 1153 tctgttgaactctattgaagaaactgacggttactcttcttcttcttcttcttcttcttctt 1212
Db 1464 ACCGTGGAGAAATACACCCAGACAGATGGATTCTCGTCTGCTTGGAGGGTTCGTTTGGT 1523
QY 1213 gctagacttcaacttgaatgatgaactgaagctgaagacacacacacacacacacacacacac 1272
Db 1524 TCGGTTTGTACGTTCGAGATGATGCAGTGTACGCGGACAGAGCGGCTGTCCGTGTC 1583
QY 1273 ttggttaacgacagagttgttccattgacggttcttcttcttcttcttcttcttcttcttct 1332
Db 1584 TTGGTTAATGATCGCTTTCGCGCTGCATGGGTGTCGCGTTGATGCTTTGGGGAGATGT 1643
QY 1333 aagagagacgacttcttgaaggttcttcttcttcttcttcttcttcttcttcttcttctt 1392
Db 1644 ACCCGGATAGCTTTTGTAGGGGTTGAGCTTTCCTAGATCTGGGGGTGATTGGGGCGGAG 1703
QY 1393 tgttctgctta 1403
Db 1704 TGTTTTCGTTA 1714

RESULT 13
LOCUS I13429 6756 bp DNA PAT 26-JUL-1995
DEFINITION Sequence 31 from patent US 5436156.
ACCESSION I13429
VERSION I13429.1 GI:910770
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6756)
AUTHORS Van Gormom,R.F.M., Van Hartingsveldt,W., Van Paridon,P.A.,
Veenstra,A.E., Luiten,R.G.M. and Seiten,G.C.M.
TITLE Cloning and expression of phytase from aspergillus
JOURNAL Patent: US 5436156-A 31 25-JUL-1995;
FEATURES
Location/Qualifiers
source 1. 6756
BASE COUNT 1615 a 1714 c 1712 g 1715 t
ORIGIN
Query Match 43.5%; Score 610.2; DB 10; Length 6756;
Best Local Similarity 64.9%; Pred. No. 4.3e-152;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;
QY 13 gtcgtactacttccatgacacacttgcgttccacatccggtacccgcttgggtcct 72
Db 324 GTGGGACTACTGATCGCTGACTATCTGTGACAGTCACTCCGGACTGGCAGTCCCGCC 383
QY 73 cgtgtaactctcaactcttctgtgacactgttgacggtgttaccatggtttcccaagaatt 132
Db 384 TCGAAGAAATCAATCCAGTTCGATACGATACGATACGATACGATACGATACGATACGAT 443
QY 133 tctcactgtggtggtacacacttccatcttcttcttcttcttcttcttcttcttcttctt 192
Db 444 TCGCATCTTTGGGGTCAATACGACACCGTCTCTCTCTGGCAACGAATCGTATCTCC 503
QY 193 ccagcgtctccaaagggtgttagaacttcttcttcttcttcttcttcttcttcttcttctt 252
Db 504 CCGTAGGTGCCCGCGGATGACAGAGTCACTTTCGCTCAGGTCTCTCCCGCTCAATGGAGCG 563
QY 253 agatacccaacttcttctgcttaagcgttactcttcttcttcttcttcttcttcttcttctt 312
Db 564 CGGTATCCGACCGACTCCCAAGGCAAGAAATACTCCGCTCTCATGTAGGAGATCCAGCAG 623
QY 313 aacgtactgttctcaagggttaagtaagcttcttcttcttcttcttcttcttcttcttctt 372
Db 624 AACGGACACCTTTGACGGAAATATGCTTCTCTGAGACATACAACTACAGCTTGGGT 683
QY 373 gctgacgacttgaactccatctcgttgaacacacaaatgttaactctgttatttaagtctac 432
Db 684 GCAGATGACCTGACTCCCTTCGGAGACAGAGTACTGTAACCTCCGGCATCAAGTTCTAC 743
QY 433 agaagatacagaaggttgggttagaagattgttccatcttcttcttcttcttcttcttctgac 492
Db 744 CAGCGGTACGAATCGCTCACAGGAACATCGTTCATTCATCCGATCCTCTGGGTCCAGC 803
QY 493 aaggtattgtctctgctgaaggttcaattgaaggttccaaatcttcaagttgggtgac 552
Db 804 CGCTGATCGCTCCGGACAGAAATTCATCGAGGGCTTCCAGAGCACCAAGCTGAAGAT 863
QY 553 ccaggtgttaaccacacacacacacacacacacacacacacacacacacacacacacacacac 612
Db 864 CCGTGTGCCAGCGCGCAATCTGCGCCCAAGATCGACGTGTCTATTCGGAGGCCAGC 923
QY 613 gtttacacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 672
Db 924 TCATCCCAACACACTCTCGACCCAGCACCTGACCTGTCTTCAAGACACAGCAATTTGGCC 983
QY 673 gacgacttgaagcttaacttcaactgtgttctgctccacacacacacacacacacacacacac 732
Db 984 GATACCGTCGAAGCAATTTACCGCCACGTTGTCGCCCTCCATTCGTCAACGTCCTGGAG 1043
QY 733 gctcaactgtccaggtgttaacttgaactgacgaagcgttgaacttgaacttgaacttgaact 792
Db 1044 AACGACCTGTCCGGTGTGACTCTCACAGACACAGAGTGAACCTACCTCATGGACATGTGC 1103
QY 793 ccattgacacactgtgttagaactctgactgaactgaactgaactgaactgaactgaactgaact 852
Db 1104 TCCTTCAGACACATCTCCACACACACCGTTCGACACCAAGCTGTCCCGCTTCTGTGACCTG 1163
QY 853 ttcactacacacacacacacacacacacacacacacacacacacacacacacacacacacacac 912
Db 1164 TTCACCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTTGAAAGATTTACGGC 1223
QY 913 tacggtgtgttaaccatgtgtccagctcaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1224 CATGGTGCAGGTAAACCGCTTCGGCCGAGCCAGCGGCTCGGCTACGCTAACGAGCTCATC 1283
QY 973 gctagatgactactctccagttcaagcaccacacttctactaaccacacacacacacacactt 1032
Db 1284 GCCCGTGTGACCCACTCGCTGTCCAGATGACACCGATTCACACGATTCACACGATTTGAC 1343
QY 1033 aaccagctacttccattgaacgctacttctgtacgctgaacttcttctcaacacacact 1092
```


QY	495	agttattgctctgctgaagaagttaattgaagggttttccaattctgctaagttggctgaccc	554
Db	591	GGTTATTGCTTCGGGAGAGAAGTTATCATGAGGGTTCCAGCAGGCGAAGCTGGCTGATCC	650
QY	555	aggTgctaaccacacacaagattctccagttatttaaacggtattattccagaaggtgctb99	614
Db	651	TGCGCGGA --- CGAACCGCGCGCTCCGGCGATTAGTGTGATTATTCGGAGACGAGAC	707
QY	615	ttacaacaacactttggaccacggtttgtgactgcttccgaagaactgaattgggtga	674
Db	708	GTTTCAACAATACGCTGGACACCGTGTGTGCACGAGAGTTTGAAGGGAGTCACTGGGAGA	767
QY	675	cgagttgaagctaacttaacttaactgctgttttccgtccaccaattagagtagattggaagc	734
Db	768	TGAGGTTGGGGCAATTTCACTTGCCTCTTTTGACCCGACATCCGAGCTCGCGCCGAGAA	827
QY	735	tcaattgcccaggtgttaacttgactgacgaagaogtltgtaacttggatggacatgtgtcc	794
Db	828	GCATCTTCCTGGCTGACCGCTGACAGACGAGGAGCGTGTCACTTAATAGACATATGTGTC	887
QY	795	attcgacactgttgtagaactctgacgtactcaattgtctccattctgtgactggtt	854
Db	888	GTTTGATACGTTAGCGCGCACCGAGCGACGAAGTCAGCTGTCAACCGTTCGTCAACTCTT	947
QY	855	cactcaacgaagaattggattcaatacagcactctgcaattcttgggtgaagtactacggtta	914
Db	948	CACCTCAATGAGTGGGAAGAAGTACAACCTACGCTTCACTCTTGGGCAAGTACTACGGCTA	1007
QY	915	cggTgcgtgtaaccatcgggtccagctcaagggtgttgggttcgtttaacgaattgattgc	974
Db	1008	CGSGCAGCAACCCCTCTGGACCGGCTCAGGGGATAGGGTTCACCAACAGAGGTGATGC	1067
QY	975	tagattgactaactctccagttcaagaccacacttctactaaccacacactttgactctaa	1034
Db	1068	CGGGTTGACTCGTTCGCCAGTGCAGGACCACACCGACACTAACTCGACTCTAGTCTCAA	1127
QY	1035	cccagctacttcccatgaaacgctacttttgacgctgaectctctccacgacacactat	1094
Db	1128	CCCGGCCACCTTCGCGTTGAACGCTACCATCTAGCTGCGACTTTTTCACACGACACAGCAT	1187
QY	1095	ggtttctatttttcgcttgggttltgtacaacggtactaaagccattgtctactacttc	1154
Db	1188	GCTTTCCATCTCTTTGCTTTGGGCTGTACACGGCACTGAACCCCTGTCCCGGACCTC	1247
QY	1155	tgttgaatctattgaagaacactgacggttactctgcttcttggactgttccattcgtgc	1214
Db	1248	GGTGTAAAAGCGCAAGTAATTGGATGGGTATTTCTGCATCTCTGGGTGGTGCCTTCGGCGC	1307
QY	1215	tagacttactgtgaaatgatgcaatgtgaagctgaaaggaaccattggttagatttt	1274
Db	1308	CGGAGCCCTACTTCGAGACGATGCAATGCAAGTCGAAAGAGGAGGCTCTTGTTCGCGCTTT	1367
QY	1275	ggttaacgacagagttgttccattgacaggttgtggtgttgacaaagtggtagatgataa	1334
Db	1368	GATTAATGACCGGGTTGTGCCACTTGCATGGCTGGGATGTGGACAAGCTGGGCGATGCA	1427
QY	1335	gagagaocacttcgtgaaggtttgtcttcgctagatctggtggttaactgggaagaatg	1394
Db	1428	GCTGAATGACTTTGTCAAGSGATTGAGTTGGGCCAGNCTGGGGGCACTGGGGAGAGTG	1487
QY	1395	tttgcgttaa 1404	
Db	1488	CTTTAGTTGA 1497	

RESULT 15

SFU59804	..	STU59804	DNA	PLN	17-JAN-1998
LOCUS		1912 bp			
DEFINITION		<i>Aspergillus fumigatus</i>	phytase gene, complete cds.		
ACCESSION		U59804			
VERSION		U59804.1	GI:2108353		
KEYWORDS					
SOURCE		<i>Aspergillus fumigatus</i> .			

QY 495 agttattgtctgtgaaagttcaattgaaggtttccaaatctcgtcaagttggtgaccc 554
 Db 832 GGTATTGCTTCGGGAGAGAAAGTTTCAGAGGGTTCAGGAGCGAAAGCTGGCTGATCC 891
 QY 555 aggtgtaaccacacccaagctctccagttattaaactattattattccagaagtgctgg 614
 Db 892 TGGCGCGGA---CGAACCGCGCGCTCCCGCGGATTAAGTGTGATTAATCCGGAGCGGAGAC 948
 QY 615 tiacacaacacttttggaccacgggttggtagtcttccgaagaatcgtgaattgggtga 674
 Db 949 GYTCAACATACGCTGGACACGGTGTGTGCACGAAGTTTGAGGCGAGTCAAGTGGGAGA 1008
 QY 675 cgaagtgaagtaacttaactgctgtgttttgcgtccacccaattagagctagatgggaagc 734
 Db 1009 TGAGGTTGGCGCAATTTCACTGCCTCTTTGACCCCGACATCCGAGCTCGCGCGAGAA 1068
 QY 735 tcaattgcccagtgtaactgaactagacgaagcgtgttaactgtatggacatgtgtcc 794
 Db 1069 GCATCTCTCGGTGACGCTCACAGACGAGGACGTTGTCAAGTCTAATGGACATGTGTTC 1128
 QY 795 attcgacactgtttagaactctgaagctactcaaatgttccattctgtgactgttt 854
 Db 1129 GTTTGATACGGTAGCGCGACCGAGCGACGAAGTCAGCTGTCAACGCTTCTGTCAACTCTT 1188
 QY 855 cactcaagcgaatggattcaatacagactactgcaatcttgggttaagtactacggtta 914
 Db 1189 CACTCAATGAGTGGAGAAAGTACAACTTACCTTGTGGGCAAGTACTACGGCTA 1248
 QY 915 cgtgtgtgtaacccattgggtccagctcaaggtgtgtgttctcgttaacgaattgttc 974
 Db 1249 CGCGCGGACCAACCCCTCTGGGACCGCTCAGGGATAGGGTTCCACCAAGAGCTGATTGC 1308
 QY 975 tagattgactactctccagttccagacacacacttctactaaccacactttggaacttaa 1034
 Db 1309 CCGGTTGACTCTGCTGCCAGTGCAGGACACACACAGCACTAAGTCTAGTCTCCAA 1368
 QY 1035 ccagctactttccattgaacgctactttgtacgtgacttctctcagacacacactat 1094
 Db 1369 CCGGCGACCTTCCGTTGAACGCTACCATGTACGTCGACTTTTCACAGGACACACAGCAT 1428
 QY 1095 ggtttctattttctcgtgttgggtttgtacaacggtactaagccattgtctactacttc 1154
 Db 1429 GGTTCATCTCTTGTGATTTGGGCTGTACACGCGACTGAACCCCTGTCCCGGACCTC 1488
 QY 1155 tgtgaatctattgaagaactgaacgttactctctctgtgactgttccattcgtctgc 1214
 Db 1489 GGTGAAAGCGCCAGGAATTTGGATGGGTATCTGATCTCTGGGTGCTGCCCTTCGGCGC 1548
 QY 1215 tagagcttaagttgaatgatgaatgtgaagctgaaaggaacccattggttagagtttt 1274
 Db 1549 CGAGGCTACTTCGAGACGATGCAATGCAAGTCGGAAGAGAGCCCTCTTGTTCGGCTTT 1608
 QY 1275 ggttaacgacagagttgttccattgacaggtgtgtgtgttgacaagttgggttagatgtaa 1334
 Db 1609 GATTAAATGACCGGTTGTGCGCACTGCTGCGATGCGATGTGGACAAAGCTGGGCGATGCAA 1668
 QY 1335 gagagacgactctgtgaaggtttgtcttctcgttagatctgtgtgtaactgggaagaatg 1394
 Db 1669 GCTCAATGACTTTGTCAAGGATTTAGTTGGGCCACATCTGGGGGCAACTGGGGAGAGTG 1728
 QY 1395 ttctgcttaa 1404
 Db 1729 CTTTAGTTGA 1738

Search completed: October 26, 2001, 17:47:25
 Job time: 8801 sec

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*
48: gb_est48:*
49: gb_est49:*
50: gb_est50:*
51: gb_est51:*
52: gb_est52:*
53: gb_est53:*
54: gb_est54:*
55: gb_est55:*
56: gb_est56:*
57: gb_est57:*
58: gb_est58:*
59: gb_est59:*
60: gb_est60:*
61: gb_est61:*
62: gb_est62:*
63: gb_est63:*
64: gb_est64:*
65: gb_est65:*
66: gb_est66:*
67: gb_est67:*
68: gb_est68:*
69: gb_est69:*
70: gb_est70:*
71: gb_est71:*
72: gb_est72:*
73: gb_est73:*
74: gb_est74:*
75: gb_est75:*
76: gb_est76:*
77: gb_est77:*
78: gb_est78:*
79: gb_est79:*
80: gb_est80:*
81: gb_est81:*
82: gb_est82:*
83: gb_est83:*
84: gb_est84:*
85: gb_est85:*
86: gb_est86:*
87: gb_est87:*
88: gb_est88:*
89: gb_est89:*
90: gb_est90:*
91: gb_est91:*
92: gb_est92:*
93: gb_est93:*
94: gb_est94:*
95: gb_est95:*
96: gb_est96:*
97: gb_est97:*
98: gb_est98:*
99: gb_est99:*
100: gb_est100:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	152.6	10.9	802	227	AQ324539	mgxb0019C
2	117.8	8.4	837	226	AQ255459	mgxb0014M
3	109	7.8	581	104	AJ274007	AJ274007
4	83.8	6.0	753	225	AQ163004	mgxb0021D
5	81.4	5.8	699	225	AQ162040	mgxb0010F
6	75.6	5.4	810	225	AQ161556	mgxb0008K
7	74.6	5.3	769	227	AQ361474	mgxb0004I
8	73.6	5.2	669	227	AQ361495	mgxb0004K
9	43.6	3.1	1016	221	CNS04ESQ	AL287459 Tetraodon
10	41.8	3.0	341	109	AV054552	AV054552
11	41.8	3.0	481	225	AQ202816	RPC111-48
12	41.8	3.0	545	224	AQ140737	AV129427
13	41.2	2.9	231	25	AV129427	AV129427
14	41.2	2.9	234	109	AV064132	AV064132
15	41.2	2.9	236	109	AV054718	AV054718
16	41.2	2.9	247	28	AV233643	AV233643
17	41.2	2.9	264	108	AV002759	AV002759
18	41.2	2.9	269	109	AV077122	AV077122
19	41.2	2.9	278	25	AV114552	AV114552
20	41.2	2.9	285	109	AV081956	AV081956
21	41.2	2.9	297	109	AV060181	AV060181
22	41.2	2.9	300	109	AV053884	AV053884
23	41.2	2.9	306	25	AV118366	AV118366
24	41.2	2.9	342	109	AV047696	AV047696
25	40.4	2.9	468	151	BF637070	BF637070
26	40.4	2.9	482	30	AV423243	AV423243
27	40.2	2.9	265	109	AV049962	AV049962
28	40.2	2.9	295	26	AV212676	AV212676
29	40.2	2.9	512	107	AQ039641	AQ039641
30	39.6	2.8	227	25	AV161297	AV161297
31	39.6	2.8	249	109	AV056003	AV056003
32	39.6	2.8	266	109	AV091633	AV091633
33	39.6	2.8	268	109	AV049564	AV049564
34	39.6	2.8	270	109	AV079904	AV079904
35	39.6	2.8	280	109	AV078019	AV078019
36	39.6	2.8	285	109	AV100697	AV100697
37	39.6	2.8	289	25	AV120593	AV120593
38	39.6	2.8	289	109	AV055109	AV055109
39	39.6	2.8	290	109	AV053894	AV053894
40	39.6	2.8	291	25	AV151459	AV151459
41	39.6	2.8	296	109	AV058485	AV058485
42	39.6	2.8	297	109	AV050799	AV050799
43	39.6	2.8	298	109	AV067677	AV067677
44	39.6	2.8	300	109	AV054136	AV054136
45	39.6	2.8	303	25	AV118481	AV118481

ALIGNMENTS

RESULT 1
 LOCUS AQ324539/c
 DEFINITION mgxb0019C01r CUG1 Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0019C01r, DNA sequence.
 ACCESSION AQ324539
 VERSION 1
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 802)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL COMMENT

Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 408.
 Location/Qualifiers
 .source
 /organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0019C01r"
 /clone_lib="CUG1 Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25X genome coverage. High density colony filters
 are available upon request."

FEATURES

.source

BASE COUNT
 ORIGIN

166 a 218 c 209 g 208 t 1 others

Query Match 10.9%; Score 152.6; DB 227; Length 802;
 Best Local Similarity 55.3%; Pred. No. 1.4e-33;
 Matches 343; Conservative 0; Mismatches 265; Indels 12; Gaps 2;

QY 744 aggtgttaactgtactgacgaagacgttggtaactgtgacatgtgctccattogacac 803
 Db 614 AGGCTGAACCTCAGCGGCAAGCAAGCGGTGCAGATCATGAGCTTCTGCCCGTTTGAGAC 555
 QY 804 tgtgttagaactctgacgtactcaattgtctcattctgtgactgttctcaacga 863
 Db 554 CGTGGCAGATCCNCAAGCCAC-----CCTGTGCGAGTCTGCACGCTGTTTACGCAAG 501
 QY 864 cgaatggttaataacgactacttgcatttgggttaagtactacggttaacggtgtg 923
 Db 500 AGACTGGGAGGCATATGACTATCTCAGACACTGGGGAAGTGGTATGGTTACGGCAATGG 441
 QY 924 taaccatttgggtccagctcaagggtgttgggttcttaacgaattgattgattgac 983
 Db 440 CAACCCCTGGGGCTCCACGCAAGGGGTGGGCTTCGTCAACGAGCTCATCGCGAGGCTGCT 381
 QY 984 tcactctcagttcaagaccacactcttactaacacacttggacttaaccagctac 1043
 Db 380 CCAAAAGCCGTTGAGACCCACACAAATACCACTGAGCTCGACAGGAGCCCATCGAC 321
 QY 1044 ttcccatgaacgtactttgtacgtgacttctctcacgacacactatggtttcat 1103
 Db 320 GTTCCCACTAGACAAAAGCTGTACGCCGACTTTAGCCATGATAACGATATGCTGGGCAT 261
 QY 1104 ttcttctgttgggtttgtacacggtactaagcattgtctactacttcttggtaac 1163
 Db 260 CTACGCCGCTGGGGATTACACGCCACGCCGCCCTTCGTTCCCAAAAGAGAG 201
 QY 1164 ta-----ttgaagaactgacggttactctgtcttctggacttctccattcgtgtag 1217
 Db 200 AAGAGCGCGCAGGAGCTCAGCGGGTTCTGTCAGCTGGGGGTACCGTTCGACGCGAG 141
 QY 1218 agcttacgttgaatgatgtgaactgtgaagctgaaaggaaccattggttagagtttgg 1277

```

Db 140 GATGTTTGTGAAAAATGACTTGGCCAGGCGAGAACGAGGAGCTTGTGAGAACTCTGGT 81
Qy 1278 taacagacaggtgttcacattgacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1337
Db 80 CAACGACAGGGTGCAGCGCTGCAGAACTGCGATGCCGACGATGGGTGCTGTCGACAGCT 21
Qy 1338 agacgacttctgtgaaggtt 1357
Db 20 GAGCAAGTCTGTTGAGAGCT 1

RESULT 2
LOCUS A0255459 837 bp DNA GSS 23-OCT-1998
DEFINITION mgxb0014M15r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
clone mgxb0014M15r, DNA sequence.
ACCESSION A0255459
VERSION A0255459.1 GI:3779774
KEYWORDS GSS.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea.
REFERENCE 1 (bases 1 to 837)
AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
Genome
JOURNAL Unpublished (1998)
CONTACT: Dean RA
Clemson University
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 60
High quality sequence stop: 277.
FEATURES
source
1. .837
/organism="Magnaporthe grisea"
/strain="70-15"
/db_xref="taxon:148305"
/clone="mgxb0014M15r"
/clone_lib="CUGI Rice Blast BAC Library"
/tissue_type="Protoplasts"
/lab_host="E. coli DH10B"
/note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."
BASE COUNT 248 a 154 c 188 g 243 t 4 others
ORIGIN

Query Match 8.4%; Score 117.8; DB 226; Length 837;
Best Local Similarity 56.2%; Pred. No. 2.2e-23;
Matches 221; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 737 acttgcaggtgttaacttgactgacgaagacgtgtgttaactgtgacatgtgtccat 796
Db 140 ATTTTCAAGGTGCTGGGTGTAAACCGTTCAGACGTTGTAAATCTATGGGACTGTGACCA 199
Qy 797 tcgacactgtgtgacactctgacgctactcaattgtctccattgtctgactgtgtca 856

```

```

Db 200 TGGAAACACAGCGCAACTTTGAAAAAATGGCCAACTTCACCGCTTTCACATCTGTTTA 259
Qy 857 ctacagacgaatgattgaatcagcactacttgcaactcttttgggtaagtactacggttacg 916
Db 260 CGAAGCAGATTGGGTAAATATGTTACTTGTCCAGCGTGCAGAAATGGTACAGATACG 319
Qy 917 gtgtgtgtaaccattgggtccagctccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 976
Db 320 GAAACGGTAATCTCTTGGGCCCACTATGCGGTGGGTGAAACGAACTTATTGCAAC 379
Qy 977 gattgactcaactctcagttcaagacacacacttctactaaccacacttggactctaac 1036
Db 380 GATTAAACCCGAAAGCCAGTCCAAAGATCAACACCATGTCATAGACACTTGCATGAACC 439
Qy 1037 cagctactttccactgaacgctactttgtacgtgactctctctcagcacaacactatgg 1096
Db 440 CGGNAACCTTACTCCTPACAGGCAAAATGTGCTGATTTATCCATACGGAGGATATTA 499
Qy 1097 ttctattttctctogcttgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1129
Db 500 TAGGGATTACGGCGCTTGTGGGTATTTAACG 532

RESULT 3
LOCUS AJ274007 581 bp mRNA EST 29-DEC-1999
DEFINITION AJ274007 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae
cDNA clone Ma#1628, mRNA sequence.
ACCESSION AJ274007
VERSION AJ274007.1 GI:6433380
KEYWORDS EST.
SOURCE Metarhizium anisopliae.
ORGANISM Metarhizium anisopliae.
REFERENCE 1 (bases 1 to 581)
AUTHORS Screen, S.E., Mathur, P. and St. Leger, R.J.
TITLE EST analysis of the insect pathogenic fungus Metarhizium anisopliae
JOURNAL Unpublished (1999)
COMMENT Contact: Screen SE
Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA.
FEATURES
source
1. .581
/organism="Metarhizium anisopliae"
/strain="ARSEF 2575"
/db_xref="taxon:5530"
/clone="Ma#1628"
/clone_lib="Metarhizium anisopliae ARSEF 2575"
/note="Vector: Unizap; Metarhizium anisopliae was grown on
insect cuticle for 24 hours. A cDNA library was
constructed in the unidirectional Lambda vector, Unizap"
BASE COUNT 138 a 153 c 146 g 144 t
ORIGIN

Query Match 7.8%; Score 109; DB 104; Length 581;
Best Local Similarity 57.8%; Pred. No. 7.4e-21;
Matches 234; Conservative 0; Mismatches 165; Indels 6; Gaps 2;

Qy 142 tgggtatcatctctccattctcttctgtgctgacgaatctgtctatttctccagagctt 201
Db 163 TGGGCCAGTACTCGCGTACTTCTC---AGCACACAGGGATCTATTATAGCCTGACATC 219
Qy 202 ccaaggggtgttagagttacttctgttcaagttttgtgttagacacaggtgtgttagatccca 261
Db 220 CCGTCGGGATGCGAGGTGACCTTGTCTCAATCTGTCTCGACATGGGTGCCGAAACCCG 279
Qy 262 acttcttctgtctcaagcgtactctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 321
Db 280 ACGGCGCGCAAGTCCAGGCCCTACAAGGACCTTGTGAGCGGATCCAAAGACGTCACG 339

```


Qy 322 gcttcaagggttaagtaagcgttcttctgaagaacttacacatacacttttggtgactgaagac 381
 Db 340 GACTATGGGAAGGGCTTCGAGTTCTCAAGATTCAAGATTACCTAGGTCTCAGGAC 399
 Qy 382 ttgaactcattcgttgaaacaaatgggttaactctggttattaaagttctacagaagatac 441
 Db 400 CTCACCCCTTGGCGGAGGAGGAATGTCAGTCTGGAAGAAAGTCTTCAAGCGCTAT 459
 Qy 442 aagcgtttggttagaa---gattgtccattcattagagcttctggttctgaagagtt 498
 Db 460 CAGAAGTACGGGAGGATTCGACCCATCCCTTTGTCGAGCCTCGGGCTCTGAAAGATC 519
 Qy 499 attgcttctgctgaagaagttcattgaaggtttccaatctgctaaag 543
 Db 520 GTCATGTCAGCGCAAGATTTTGTTCATGCTTCTACAAAGCCAAAG 564

RESULT 4
 LOCUS AQ163004 753 bp DNA GSS 09-SEP-1998
 DEFINITION mgxb0021D19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0021D19r, DNA sequence.
 ACCESSION AQ163004
 VERSION AQ163004.1 GI:3559405
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 753)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinoski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 448.
 Location/Qualifiers
 1..753
 /organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0021D19r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."

BASE COUNT 208 a 180 c 156 g 206 t 3 others
 ORIGIN

Query Match 6.0%; Score 83.8; DB 225; Length 753;
 Best Local Similarity 50.7%; Pred. No. 2e-13;
 Matches 227; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

Qy 90 ttgtgacactgttgacggttggtttaccatgtttccacagaaatttctcacttgggtgac 149
 Db 189 TTTTGATGCGCAGCCCGCGGTTCATTTAAAGCAGCAATCACTCATATTTGGGCCA 248
 Qy 150 ataactcattctctctcttctggtgacgaatctgctatttctcagacgttccaa---- 205
 Db 249 GTATGCGACCATTTTCGTTGACCGTCAGCACCATCGAATATGATTCCTCTGTACCTT 308
 Qy 206 --aggggttagagttacttctcgttcaagtttcttagacacggtgctagatacccaac 263
 Db 309 GCCAGGTTGTAGCATTTAGCTTTGCCAAGTCTTCAGAGCAGCGGGCTCATACCCAAC 368
 Qy 264 tcttctgctgcttaagcgctactctgcttggattgaagctattcaaaagaacgctactgc 323
 Db 369 CGCTCAAAACCGCGCGAAATTCGAAACACGATTTCACCGACTCCAAACCTCACGTACCG 428
 Qy 324 ttcaagggttaagtacgcttcttgaagacttacacatacactttgggtgctgacgactt 383
 Db 429 TAGTGGCGTATTGAATACATACATTAATAAATACTCAATACTCGAGTCGAGGAATN 488
 Qy 384 gaetccattcgttgaaacaaatgggttaactctggtattaaagttctacagaagatacaa 443
 Db 489 GAATGATTTGGCGCTCGGCAAAACAGAGAACTCGGGTATTNATTTTACCAGAGTACCA 548
 Qy 444 ggcttggctagaagaattgtccattcattagagcttctggttctgacagagttattgc 503
 Db 549 AAATCTGCCGAGAGGAACGACCATTTATTCGTATGATGACAAANAACCGCTGTTGA 608
 Qy 504 ttctgctgaaagtctcattgaggttttc 531
 Db 609 TAGCGCAGAACTCTGGGCTAGGGGTTTC 636

RESULT 5
 LOCUS AQ162040/c

DEFINITION AQ162040 699 bp DNA GSS 09-SEP-1998
 clone mgxb0010F06r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0010F06r, DNA sequence.
 ACCESSION AQ162040
 VERSION AQ162040.1 GI:3558441
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea.
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 699)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinoski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished (1998)
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 285.
 Location/Qualifiers
 1..699
 /organism="Magnaporthe grisea"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0010F06r"
 /clone_lib="CUGI Rice Blast BAC Library"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."

FEATURES
 Source

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

[illegible]

RESULT	6
AQ161556/c	
LOCUS	810 bp DNA GSS 09-SEP-1998
DEFINITION	mgxb00080r2r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb00080r2r, DNA sequence.
ACCESSION	AQ161556
VERSION	AQ161556.1 GI:3557957
KEYWORDS	GSS.
SOURCE	Magnaporthe grisea.
ORGANISM	Magnaporthe grisea
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina: Sordariomycetes; Sordariomycetes:incertae sedis; Magnaporthaceae; Magnaporthae. 1 (bases 1 to 810) Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A. A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
TITLE	AUTHORS

FEATURES	source
high quality reads	1. 810
Location/Qualifiers	/organism="Magnaporthe grisea"
	/strain="70-15"
	/db_xref="taxon:148305"
	/clone="mqxb0008K02r"
	/clone_lib="CGI Rice Blast BAC Library"
	/tissue_type="Protoplasts"
	/lab_host="E. coli DH10B"
	/note="vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request." 196 a 219 c 188 q 207 t

BASE COUNT	196 a	219 c	188 g	207 t	
ORIGIN					
Query Match	5.4*	Score 75.6;	DB 225;	Length 810;	
Best Local Similarity	58.4*;	Pred. No. 5.1e-11;			
Matches 132; Conservative	0;	Mismatches 94;	Indels	0;	Gaps
QY	1179	cggttaactctgctcttgagactctgccactcgctctgctagagcttcagttgaaatgatca	1238		
Db	650	CGGGTTCGTCGTCACGTCGGCGGTACCGTTCGCAGCGCAGGATGTTGTTGTAATAATGAC	591		
QY	1239	atgtgaagctgaaaaaggaacattggtttagattggttttaacacagagttgttccatt	1298		
Db	590	TTGCGCAGGGGAGAACGACGAGCAGCTGTGTGAATACCTCGTCAACGACAGCGGTGACCCGCT	531		
QY	1299	gcacggtgtggtgttaacagttggttatgtatgaagagagacgacttcgttgaaggttt	1358		
Db	530	GCAGAAATCCGATGCCGACAGATATGGGTGCTTGCACGCTGACCAAGTTCGTTGAGAGCTT	471		
QY	1359	gtctttcgtagatcgtgtgtaactgggaagaatgttttcgcttaa	1404		
Db	470	AAGCTCTCGGAGGAGTGGAGGTGCTCGGATCAATGTTTGTCTGA	425		

[illegible]

FEATURES	SOURCE
Location/Qualifiers	
1. .769	
/organism="Magnaporthe grisea"	
/strain="70-15"	
/db_xref="taxon:149305"	
/clone="mgx000411zf"	
/clone_lib="CUGI Rice Blast BAC Library"	
/tissue_type="Protoplasts"	
/lab_host="E. coli DH10B"	
/note="vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases"	

QY	1033	aaccagcgtactttcccaattgaa	1055
Db	27	AATCCGAAACAGCTTCCCGTTGA	5
RESULT 9			
CNS04ESQ	LOCUS	CNS04ESQ	1016 bp DNA GSS 21-MAY-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 105015 of library G from Tetraodon nigroviridis, genomic survey sequence.		
ACCESSION	AL287459	GI:8025953	
VERSION	AL287459.1	GI:8025953	
KEYWORDS	GSS; genome survey sequence.		
SOURCE	Tetraodon nigroviridis.		
ORGANISM	Tetraodon nigroviridis.		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.			
REFERENCE	1 (bases 1 to 1016)		
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot.A. and Weissenbach,J.		
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1016)		
TITLE	Saurin,W. and Weissenbach,J.		
JOURNAL	Bernot.A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin.W. and Weissenbach,J.		
AUTHORS	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence		
TITLE	Unpublished		
REFERENCE	3 (bases 1 to 1016)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases		
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.		
FEATURES			
Source	Location/Qualifiers		
1..1016	/organism="Tetraodon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="105015"		
	/clone_lib="G"		
	/note="Genoscope sequence ID : C0BGL05AH08UP1-end : T7"		
BASE COUNT	227 a 259 c 231 g 287 t 12 others		
ORIGIN			
Query Match	3.1%; Score 43.6;	DB 221;	Length 1016;
Best Local Similarity	48.0%;	Pred. No. 0.13;	
Matches	122; Conservative	1; Mismatches 131;	Indels 0; Gaps 0;
QY	575	cttctccagtattaacgttattatccagaaggctgctgttacacaacacctttggacc	634
Db	722	CTGCTGTGTACTGCTGTGTACTGCTGTGTACTGCTGTATTWACTGCTGTGTA	781
QY	635	acgggttgtagctgttcctogaagaatcgattgggtgaacagctgaagctaactca	694
Db	782	CTGCTGTGTACTGCTGTGTACTGCTGTGTACTGCTGTATTWACTGCTGTGTA	841
QY	695	ctgctgtttgcgtccaccaattagactagattggaagctcactgcagctgttaact	754
Db	842	CTGCTGTGTGTHACTGATGTTATTACTGCTGTGTTACTGCTGTGTTATTATA	901
QY	755	tgactgcagaagcgttgtttaactgatggacatgtgtccactgcacacctgttgtagaa	814
b	902	CTGCTGTGTACTGCTGTGTGTACGGCTGTGTTACTGCTGTTATTACTGCTGTGTA	961

```

RESULT 11
LOCUS AQ202816
DEFINITION RPCI11-48J14.TK RPCI-11 Homo sapiens genomic clone RPCI-11-48J14,
DNA sequence.
ACCESSION AQ202816
VERSION AQ202816.1 GI:3613607
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 481)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
JOURNAL
TITLE Other_GSSs: RPCI11-48J14.TJ
AUTHORS
CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.
FEATURES
Location/Qualifiers
1..481
/organism="Homo sapiens"
/db_xref="GDB:7518277"
/db_xref="taxon:9606"
/clone="RPCI-11-48J14"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 150 a 99 c 81 g 151 t
ORIGIN

Query Match 3.0%; Score 41.8; DB 225; Length 481;
Best Local Similarity 45.5%; Pred. No. 0.35;
Matches 148; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 966 attgattgtagatgactcactctccagttcaagaccacacttctactaacacacacttt 1025
|||||
Db 58 ATTTATTATGCTTGAATTTCTCGAGACTACTGATTTCTGCTCCATCCCACTCT 117
|||||

QY 1026 ggaactcaaccagctactttcccaattgaacgactactttgtaacgtgactctctcaaga 1085
|||||
Db 118 ATACACCTCCCTCCACTTCCAAATTAAGTGGTAATCTAGCACCTTACTGAATCTTCC 177
|||||

QY 1086 caacactagggttctattttctgcgtttgggttggttgatacaacggtactaagccattgtc 1145
|||||
Db 178 AGATAGAGAGCTCTTAAGATGGTCCACCCCGGCTTTTACACACCATATCAGAAAAAGTTCG 237
|||||

QY 1146 tactacttctgttaactattgaagaaactgacggttactctgtctcttctgactgttcc 1205
|||||
Db 238 TTATATTACACTGAAGCTCTCGAAATGAGGATGATGAAGTTCAAGATGACTTTTAC 297
|||||

QY 1206 attcgtcgttagagcttactggttgaaatgatgcaatgtgaagctgaaaggaaccattggt 1265
|||||
Db 298 CTGCGTACCTTTTGTATACCTTGGCTAAGAAACTAGTCAGACGTTTGAAAAACAAGTA 357
|||||

QY 1266 tagagtttgggttaacgacagatt 1290
|||||

Db 358 AAAGGTATCTTTAGTAAAGATT 382

RESULT 12
LOCUS AQ140737/c
DEFINITION HS_3124_A2_D10_T7 CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3124 Col=20 Row=G, DNA sequence.
ACCESSION AQ140737
VERSION AQ140737.1 GI:3531390
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 545)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
99380589
CONTACT: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3124 row: G Column: 20
Class: BAC ends
High quality sequence stop: 545.
Location/Qualifiers
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate=3124 Col=20 Row=G"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 185 a 90 c 90 g 178 t 2 others
ORIGIN

Query Match 3.0%; Score 41.8; DB 224; Length 545;
Best Local Similarity 54.0%; Pred. No. 0.36;
Matches 129; Conservative 0; Mismatches 107; Indels 3; Gaps 2;

QY 284 actctgcttggattgaagctattcaaaagacgctactcttcaagggttaagctctt 343
|||||
Db 474 AATCTGATTGGTTTTTAATAGTAAAGTACACTGGTGGAGACATGAGTACGATCTT 415
|||||

QY 344 tcttgaagacttacaactacactttgggtgctgacgactgactccat-tcggtaacaa 402
|||||
Db 414 TATTCATATATATCTCTTAATATGTGTATGACTATTTATCATCATGACAGCTCAG 355
|||||

QY 403 caaatggttaactctggttatttaagttcacagaatacagaagcttggctgagaagatt 462
|||||
Db 354 TACAAGGTTAAGTAAACCATTTATTTCTATCTGTGAAACAAG--AATAACTAGATATTTA 297
|||||

QY 463 gttccattcattagagcttctggttctgacagagttattgctctgctgaaagttcat 521
|||||
Db 296 AAGCTATACATTTAGGCTTTCTGTTCTTCAAGGGTTATATCTTCTTCTAGGCTTCAT 238
|||||

RESULT 13
LOCUS AV129427
DEFINITION AV129427 Mus musculus c57BL/6J 11-day embryo Mus musculus cdNA
clone 2700068B02, mRNA sequence.

```


Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H., Owa, C.,
Sato, K., Shibata, Y., Shigenoto, Y., Shiraki, T., Sogabe, I., Sugahara
Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tominaga, N.,
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,
Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genomeres@riken.go.jp

Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.

FEATURES
source

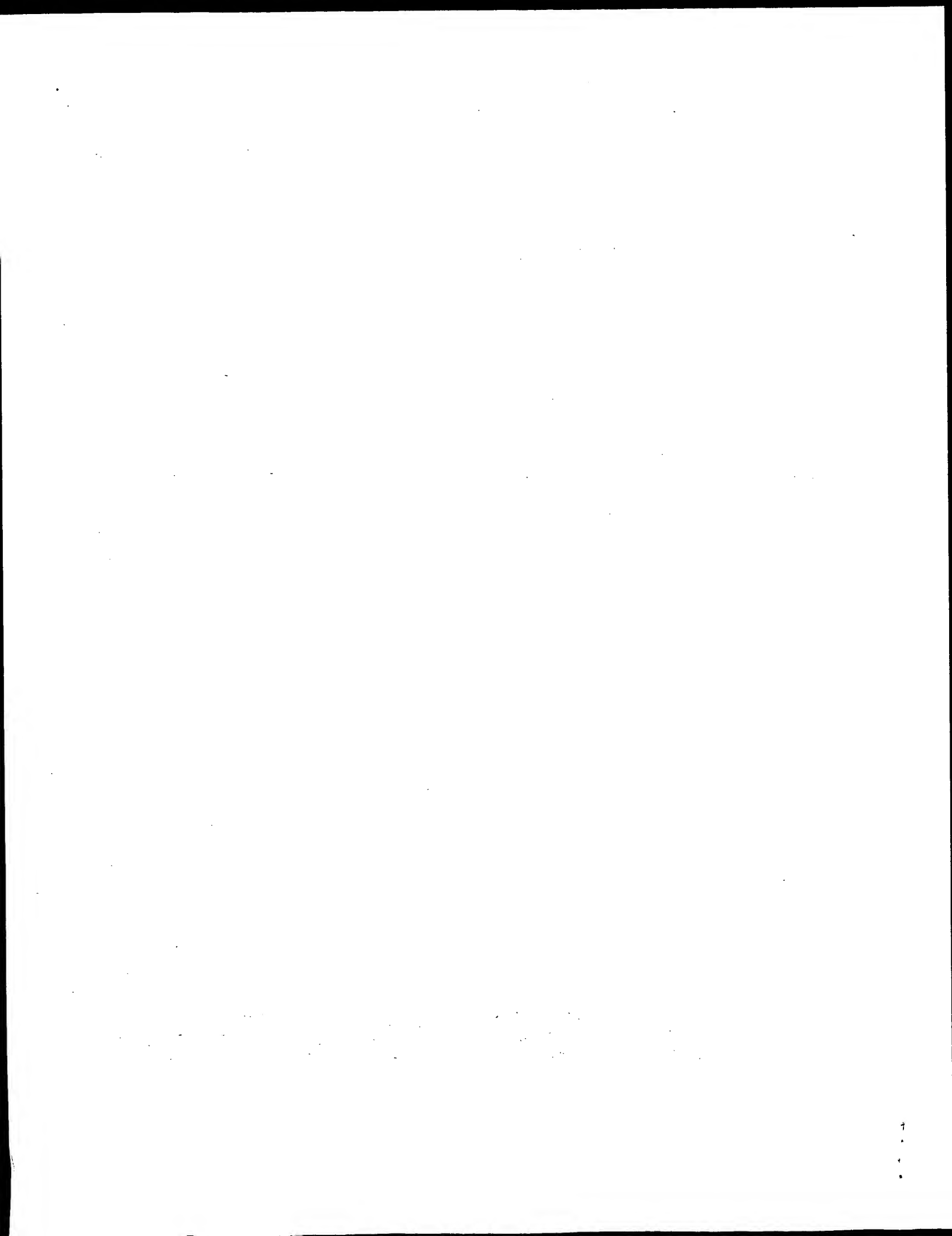
Location/Qualifiers
1..236

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1810032D19"
/clone_lib="Mus musculus pancreas C57BL/6J adult"
/sex="male"
/tissue_type="pancreas"
/dev_stage="adult"

BASE COUNT 56 a 47 C 69 g 64 t
ORIGIN

Query Match 2.9%; Score 41.2; DB 109; Length 236;
Best Local Similarity 52.3%; Pred. NO. 0.42;
Matches 91; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 1154 ctgttgaaatctattgaagaactgacggttactctgtcttggactgttcattcgctg 1213
DB 45 CTGGTGGGCCCGCTCCAGCAGCTGGAGCTGGCGCTGCTGGTACTGCTTCATCCACTG 104
QY 1214 ctgagacttacgttgaatgatgaatgaagctgaagctgaagaagaccattggttagagttt 1273
DB 105 CCGCGCCCCCAGCTGAGGAGAGAAAGTGGAGGCAAGAGAGAGATTCGAGGATTTTG 164
QY 1274 tggtaacgacagagttgttccattgcacggttgtgtgttgacaagtgggta 1327
DB 165 AAGATGACATGGGCTTCGTTTTTTTGTGACTAACTGTTTTTTTGAAGTTGGGTA 218

Search completed: October 26, 2001, 18:33:06
Job time: 6796 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 17:50:32 ; Search time 210.38 Seconds
(without alignments)
4190.388 Million cell updates/sec

Title: US-09-488-265-30

Perfect score: 1404

Sequence: 1 atggcggttcgttcgtct.....gggaagaatttctgctaa 1404

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1:	/SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1404	100.0	1404	20	AAZ31521
2	1404	100.0	1404	21	AAZ3234
3	1404	100.0	1404	21	AAZ59716
4	1388	98.9	1404	21	AAZ73293
5	1388	98.9	1426	21	AAZ73232
6	1388	98.9	1426	21	AAZ59642
7	1384.8	98.6	1404	21	AAZ73292
8	1327.2	94.5	1404	20	AAZ31520
9	1327.2	94.5	1404	21	AAZ73233
10	1327.2	94.5	1404	21	AAZ59715
11	1322.4	94.2	1404	21	AAZ73291
					Consensus phytase
					Consensus phytase
					Consensus phytase
					DNA encoding a mut
					Consensus phytase
					Consensus phytase
					DNA encoding phyta
					Consensus phytase
					Consensus phytase
					Consensus phytase
					DNA encoding a mut
					Consensus phytase

12	1319.2	94.0	1404	21	AAZ73290	Consensus phytase
13	1317.6	93.8	1426	20	AAZ27423	Ascomycete Consens
14	1317.6	93.8	1426	20	AAZ23022	Fungal phytase gen
15	1317.6	93.8	1426	21	AAZ73231	Consensus phytase
16	1317.6	93.8	1426	21	AAZ59637	DNA encoding phyta
17	1258.4	89.6	1426	20	AAZ31523	Consensus phytase
18	1258.4	89.6	1426	21	AAZ73236	Consensus phytase
19	1258.4	89.6	1426	21	AAZ59738	DNA encoding phyta
20	653.4	46.5	1404	20	AAZ31522	A. fumigatus phyta
21	653.4	46.5	1404	21	AAZ73235	Aspergillus fumiga
22	653.4	46.5	1404	21	AAZ59717	DNA encoding a mut
23	615.8	43.9	1404	18	AAZ65136	Aspergillus niger
24	614.2	43.7	1404	12	AAQ11175	Aspergillus niger
25	614.2	43.7	1404	20	AAZ27421	Chromosomal phytas
26	611	43.5	1404	12	AAQ13878	Aspergillus ficum
27	610.2	43.5	1404	15	AAQ58126	Phytase gene, Asp
28	610.2	43.5	1404	15	AAQ56944	Phytase gene, Asp
29	610.2	43.5	1404	12	AAQ11174	A. niger phytase g
30	610.2	43.5	1404	18	AAZ65137	Sequence, from ove
31	610	43.4	1571	19	AAZ03144	Aspergillus ficum
32	610	43.4	1571	20	AAZ27422	Aspergillus niger
33	583.6	41.6	1515	18	AAZ03142	Aspergillus niger
34	576.8	41.1	1931	19	AAZ03142	Aspergillus niger
35	576.8	41.1	1931	20	AAZ27413	A. nidulans phytas
36	558	39.7	1912	17	AAZ03743	Phytase chimeric g
37	539.4	38.4	1845	19	AAZ03143	Talaromyces thermo
38	537.8	38.3	1845	20	AAZ27416	T. thermophilus ph
39	528.6	37.6	2327	17	AAZ03736	Phytase gene, Asp
40	528.6	37.6	2327	20	AAZ27414	A. terreus phytase
41	527	37.5	1567	19	AAZ03145	Aspergillus terreu
42	511.2	36.4	1922	22	AAZ77028	P. hordei phytase D
43	510.6	36.4	1464	22	AAZ77030	P. hordei phytase c
44	510.6	36.4	1584	22	AAZ77029	Part of P. hordei p
45	414.2	29.5	4898	22	AAZ77031	P. hordei phytase c

ALIGNMENTS

RESULT 1

AAZ31521
ID AAZ31521 standard; DNA; 1404 BP.

AC AAZ31521;

XX 06-JAN-2000 (first entry)

XX Consensus phytase-10-thermo(3)-Q50T-K91A coding sequence.

XX Phytase; animal feed preparation; thermostable phytase; transgenic plant;

XX consensu sequence; ds.

OS Synthetic.

XX WO9948380-Al.

XX 30-SEP-1999.

XX 22-MAR-1999; 99WO-DK00154.

XX 23-MAR-1998; 98DK-0000407.

XX 19-JUN-1998; 98DK-0000806.

XX 18-SEP-1998; 98DK-0001176.

XX 22-JAN-1999; 99DK-0000091.

XX 22-JAN-1999; 99DK-0000093.

XX (NOVO) NOVO-NORDISK AS.

XX Petersen S;

XX WPI; 1999-591030/50.

XX P-PSDB; AAY43170.

preparing animal feed using a thermostable phytase

Example 3: Fig 10: 71 pp: English.

This sequence encodes the consensus phytase-10-thermo(3)-050T-K91A.

This sequence encodes the consensus phytase in thermophilic *Clostridium*. This invention relates to a process for preparing animal feed comprising agglomerating feed ingredients with a thermostable phytase which is added before or during agglomeration. These plants are useful in the feed for expelling animal feed itself. The thermostable phytase allows animal feed to be produced more efficiently, in addition to improved phytase-expressing transgenic plants, these plants provide a feed ingredient and a feed additive (phytase) simultaneously.

commence 1404 BP. 320 A. 300 C. 308 C. 458 T. 0 other.

```

100.0%      Score 1404:   DB 20: Length 1404:
          rrv Match

```

	Local Similarity	100.0%;	Pred. NO. 0;	Mismatches	0;	Indels	0;	Gaps
	1404;	Conservative						
1	atggcggtgtctgctgctactgtccattgccactgtgtcggttccacatccggtacc	60						
1	atggcggtgtctgctgctactgtccattgccactgtgtcggttccacatccggtacc	60						
61	gccttgggtctcgttggttaactctcactcttcttgacactgttgagcgtgttaccaatgt	120						
61	gccttgggtctcgttggttaactctcactcttcttgacactgttgagcgtgttaccaatgt	120						
121	ttccacagaaattctcaactgttggttgacatactctccacttctctcttggctgacgaa	180						
121	ttccacagaaattctcaactgttggttgacatactctccacttctctcttggctgacgaa	180						
181	tctgctattctccagacgttccaaagggtgtgagagttacttctgataagtttctgt	240						
181	tctgctattctccagacgttccaaagggtgtgagagttacttctgataagtttctgt	240						
241	agacacggtctgatatccacaaacttcttgcgtctctaaagggtactctgctttgattga	300						
241	agacacggtctgatatccacaaacttcttgcgtctctaaagggtactctgctttgattga	300						
301	gctattcaaaagaacgtactgcttccaagggttaagtacgcttcttgaaagacttacaac	360						
301	gctattcaaaagaacgtactgcttccaagggttaagtacgcttcttgaaagacttacaac	360						
361	tacactttgggtgctgacgacttgactccattcgggtgaacacaaatggttaactctggt	420						
361	tacactttgggtgctgacgacttgactccattcgggtgaacacaaatggttaactctggt	420						
421	attaagttctacagaatatacaaggctttggctagaagaattttccattcattagagct	480						
421	attaagttctacagaatatacaaggctttggctagaagaattttccattcattagagct	480						
481	tctggttctgacagattattgctctgctgaaagtctaatgaaagtttccaactctgct	540						
481	tctggttctgacagattattgctctgctgaaagtctaatgaaagtttccaactctgct	540						
541	aagtttggtgacctgactgatacccaaccaagcttctccagttattaaagttattatt	600						
541	aagtttggtgacctgactgatacccaaccaagcttctccagttattaaagttattatt	600						
601	ccagaaggtgctggtttacaacaaactttggaccaaggtttgtgactgctttcgaaga	660						
601	ccagaaggtgctggtttacaacaaactttggaccaaggtttgtgactgctttcgaaga	660						
661	tctgaattgggtgacgacttgaaagcttaacttcaactgctgttttgcgtccaccaattga	720						
661	tctgaattgggtgacgacttgaaagcttaacttcaactgctgttttgcgtccaccaattga	720						
721	gctagattggaagctcaacttgcagggtgttaacttgactgacgaagacgttgttaacttg	780						
721	gctagattggaagctcaacttgcagggtgttaacttgactgacgaagacgttgttaacttg	780						

781	atggacatgtgtccatttcgacactgtttgctagaaacttctgacgctactcaattgtgtctcca	840
Qy		
781	atggacatgtgtccatttcgacactgtttgctagaaacttctgacgctactcaattgtgtctcca	840
Db		
841	ttctgtgacttgttcactcacgacgaaatgattcaatacagactacttgcgaattcttgggt	900
Qy		
841	ttctgtgacttgttcactcacgacgaaatgattcaatacagactacttgcgaattcttgggt	900
Db		
901	aagtactacggttacggtgtctgtgtaaacccatttgggtccagctcaaggtgtgtgttcggt	960
Qy		
901	aagtactacggttacggtgtctgtgtaaacccatttgggtccagctcaaggtgtgtgttcggt	960
Db		
961	aacgaattgattgttagattgactcaactctccagttcaagaccacacttctactaaccac	1020
Qy		
961	aacgaattgattgttagattgactcaactctccagttcaagaccacacttctactaaccac	1020
Db		
1021	aatttggactctaaacccagctactttcccatgaaacgctactttgtacgctgacttctct	1080
Qy		
1021	aatttggactctaaacccagctactttcccatgaaacgctactttgtacgctgacttctct	1080
Db		
1081	cacgacaaacactatggtttctattttcttcgcttttgggtttgtacaacggtactaangcca	1140
Qy		
1081	cacgacaaacactatggtttctattttcttcgcttttgggtttgtacaacggtactaangcca	1140
Db		
1141	ttgtctactactcttctgtgaactattatgaagaacactgacggttactctgcttcttggaact	1200
Qy		
1141	ttgtctactactcttctgtgaactattatgaagaacactgacggttactctgcttcttggaact	1200
Db		
1201	gttccatttcgtctagagcttaagttgaaatgatgatgaatgtgaagctgaaaggaaacca	1260
Qy		
1201	gttccatttcgtctagagcttaagttgaaatgatgatgaatgtgaagctgaaaggaaacca	1260
Db		
1261	ttgggttagagttttgggttaacgacagagtgttccatttcgacggttctgggtgtgacaag	1320
Qy		
1261	ttgggttagagttttgggttaacgacagagtgttccatttcgacggttctgggtgtgacaag	1320
Db		
1321	ttgggttagatgtaaagagagacgacttcttgaaggtttcttcttcgactaactctgggtggt	1380
Qy		
1321	ttgggttagatgtaaagagagacgacttcttgaaggtttcttcttcgactaactctgggtggt	1380
Db		
1381	aactgggaagaatgttttcgcttaa	1404
Qy		
1381	aactgggaagaatgttttcgcttaa	1404
Db		

RESULT 2
AA273234

AAA/3234
ID AAA73234 standard: DNA: 1404 BP.

XX
AC AAA73234:XX
DM
05-DEC-2000 (first entry)

XX

XX
KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
KW temperature stability; pH profile; temperature profile; reaction rate;
KW specific activity; substrate specificity; substrate cleavage pattern;
KW substrate binding; position specificity; phytate degradation rate;
KW food; feed; phytate; manure; ds.

XX
SyntheticXX
EW
W030043503-21

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

XX

XX

PR 22-JAN-1999; 99DK-0000032.
PR 21-SEP-1999; 99DK-0001340.

XX
PA
(NOVO) NOVO NORDISK AS.

PA

OS Aspergillus niger var. awamori.
OS Aspergillus niger str. NRRL3135.
OS Aspergillus fumigatus ATCC13073.
OS Aspergillus fumigatus ATCC32722.
OS Aspergillus fumigatus ATCC58128.
OS Aspergillus fumigatus ATCC26906.
OS Aspergillus fumigatus ATCC32239.
OS *Emicella nidulans*.
OS *Talaromyces thermophilus* ATCC20186.
OS *Myceliophthora thermophila*.
OS *Paxillus involutus* NN005693.
OS *Trametes pubescens* NN9343.
OS *Agrocybe pediades* NN009289.
OS *Peniophora lycii* NN006113.
OS *Thermomyces lanuginosa*.
OS Synthetic.

XX Key Location/Qualifiers
FT 1. 1404
FT /*tag= a
FT /product= "phytase-10-thermo[3-Q50T-K91A"]

XX EP969089-A1.
XX 05-JAN-2000.
XX 23-JUN-1999; 98EP-0111949.
XX 29-JUN-1998; 98EP-0111960.
XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX Brugger R, Lehmann M, Wyss M;
XX WPI; 2000-099429/09.
XX New stabilized enzyme formulation, useful for feed compositions for
XX monogastric animals -
XX Example 5; Fig 19; 10lpp; English.

XX The invention relates to a novel stabilised dry or liquid enzyme
XX formulation, comprising phytase (myo-inositol hexakisphosphate
XX phosphohydrolase) and one or more stabilising agents including
XX xylitol or ribitol; polyethylene glycols with a molecular weight of 600
XX to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
XX glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
XX The stabilised phytase formulation is used in a method for preparing a
XX feed composition for monogastric animals (e.g., pigs, poultry) and
XX provides a monogastric animal with its dietary requirements of
XX phosphorus. Although a large amount of phosphate is present in animal
XX feed in the form of phytate phosphorus, monogastric animals are unable
XX to utilise this form of phosphate, resulting in the addition of extra
XX phosphate to the feed of such animals. Phytase enhances the nutritional
XX value of plant material without the need for adding additional phosphate
XX to the feed. The level of phosphate pollution in the environment is
XX reduced by adding phytase to animal feed, as the animal can make use of
XX the inorganic phosphate liberated from phytate phosphorus using the
XX enzyme. The phytase formulation of the invention has an improved
XX thermostability and can therefore remain stable during long-term storage
XX and can withstand feed processing methods such as extrusion, expansion
XX and pelleting. The present sequence represents DNA encoding a mutant
XX phytase-10 consensus sequence, phytase-10-thermo[3]-Q50T-K91A, which has
XX a temperature optimum and melting point 4 degrees Celsius higher than
XX that of phytase-10 (RAY69566). Its specific activity with phytate as a
XX substrate is also strongly increased.

XX Sequence 1404 BP; 329 A; 309 C; 308 G; 458 T; 0 other;
XX

Query Match 100.0%; Score 1404; DB 21; Length 1404;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggggcggtgttcgtgctactgtccattgccacctgttgcgttccacatccgggtacc 60
DB 1 atggggcggtgttcgtgctactgtccattgccacctgttgcgttccacatccgggtacc 60
QY 61 gcttgggttcctcggtgtaactctcactctctgtgacactgttgacaggtggttaccatgt 120
DB 61 gcttgggttcctcggtgtaactctcactctctgtgacactgttgacaggtggttaccatgt 120
QY 121 ttcccaaaaatttccactgttgggttacatctcactctcattctcttcttggctgaacaa 180
DB 121 ttcccaaaaatttccactgttgggttacatctcactctcattctcttcttggctgaacaa 180
QY 181 tctgctatttctccagacgttccaaaagggtgtgaggttacttctcgttccaaagtgtgtct 240
DB 181 tctgctatttctccagacgttccaaaagggtgtgaggttacttctcgttccaaagtgtgtct 240
QY 241 agacacgggtgctagatatacccaactctctcgtgttaaggcgtactctgctttgattgaa 300
DB 241 agacacgggtgctagatatacccaactctctcgtgttaaggcgtactctgctttgattgaa 300
QY 301 gctattcaaaaagacgctactgctttcaagggttaagtacgctttcttgaagacttacaac 360
DB 301 gctattcaaaaagacgctactgctttcaagggttaagtacgctttcttgaagacttacaac 360
QY 361 tacactttgggtgctgacgacttgactcctcattcgttgaaacaaatggttaactctggt 420
DB 361 tacactttgggtgctgacgacttgactcctcattcgttgaaacaaatggttaactctggt 420
QY 421 attaatctacagaagatacaaggcttggcttgagaaagattgttccattcattagagct 480
DB 421 attaatctacagaagatacaaggcttggcttgagaaagattgttccattcattagagct 480
QY 481 tctggttctgacagagttattgcttctgctgaaaagtctcattgaaaggtttccaactgct 540
DB 481 tctggttctgacagagttattgcttctgctgaaaagtctcattgaaaggtttccaactgct 540
QY 541 aagtggctgacccagggtgtaaccacacacacacacacacacacacacacacacacacac 600
DB 541 aagtggctgacccagggtgtaaccacacacacacacacacacacacacacacacacacac 600
QY 601 ccagaaaggtgctggttacaac 660
DB 601 ccagaaaggtgctggttacaac 660
QY 661 tctgaattgggtgacgacgcttgaagctaaactcactgctgttttgcgtccacaaataga 720
DB 661 tctgaattgggtgacgacgcttgaagctaaactcactgctgttttgcgtccacaaataga 720
QY 721 gctagattggaagctcacttgcacaggtgttaacttgactgacgaagacgtgtgtaacttg 780
DB 721 gctagattggaagctcacttgcacaggtgttaacttgactgacgaagacgtgtgtaacttg 780
QY 781 atgacacatggtccattcgacactgttgtagaacttctgacgtactcactcaattgtctcca 840
DB 781 atgacacatggtccattcgacactgttgtagaacttctgacgtactcactcaattgtctcca 840
QY 841 tctgtgacttgttctcactcagcgaatgatttcaatcagactacttgcgaacttcttgggt 900
DB 841 tctgtgacttgttctcactcagcgaatgatttcaatcagactacttgcgaacttcttgggt 900
QY 901 aagtactacggttaccgtgtgtgttaacccattgggtccagctcaaggtgtgttctggt 960
DB 901 aagtactacggttaccgtgtgtgttaacccattgggtccagctcaaggtgtgttctggt 960
QY 961 aacgaattgattgttagattgactcactctcacttccagttcaagaccacacttctactaac 1020
DB 961 aacgaattgattgttagattgactcactctcacttccagttcaagaccacacttctactaac 1020
QY 1021 actttggacttaaccacagctacttctccattgaacgtacttctgttgcgtgacttctct 1080
DB 1021 actttggacttaaccacagctacttctccattgaacgtacttctgttgcgtgacttctct 1080

CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence encodes a phytase sequence from the present
 CC invention.
 XX
 SQ Sequence 1404 BP; 331 A; 311 C; 303 G; 459 T; 0 other;

Query Match 98.6%; Score 1384.8; DB 21; Length 1404;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1392; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 atggcggtgttcgtgctactgtccattgcccacattgttgggttccacatccgtacc 60
 DB 1 atggcggtgttcgtgctactgtccattgcccacattgttgggttccacatccgtacc 60
 QY 61 gcttgggtccctcggtgaattctcactctgtgacactgttgacactgttgacactgtt 120
 DB 61 gcttgggtccctcggtgaattctcactctgtgacactgttgacactgttgacactgtt 120
 QY 121 ttcccaagaatttctcactgttgggttacatctctccattctctcttcttggctgacgaa 180
 DB 121 ttcccaagaatttctcactgttgggttacatctctccattctctcttcttggctgacgaa 180
 QY 181 tctgtatttctccagacgttccaaagggtgttagatttacttcttccaaagtgttctct 240
 DB 181 tctgtatttctccagacgttccaaagggtgttagatttacttcttccaaagtgttctct 240
 QY 241 agacacggtgttagatatacccaacttcttctgctgcttaaggcgttaactctgttattgaa 300
 DB 241 agacacggtgttagatatacccaacttcttctgctgcttaaggcgttactctgttattgaa 300
 QY 301 gctattcaaaagaacgtactactcttccaaagggttaagtaagcttcttcttggctgacgaa 360
 DB 301 gctattcaaaagaacgtactactcttccaaagggttaagtaagcttcttcttggctgacgaa 360
 QY 361 tacacttgggtgtgacgactgtactcttccaaagggttgaacacaaatgttactctgtt 420
 DB 361 tacacttgggtgtgacgactgtactcttccaaagggttgaacacaaatgttactctgtt 420
 QY 421 attagttctacagaagatatacagaagcttcttctgctgcttaaggcgttactctattagact 480
 DB 421 attagttctacagaagatatacagaagcttcttctgctgcttaaggcgttactctattagact 480
 QY 481 tctggttctgacagagtattcttctgctgctgctgctgctgctgctgctgctgctgctgct 540
 DB 481 tctggttctgacagagtattcttctgctgctgctgctgctgctgctgctgctgctgctgct 540
 QY 541 aagttggtgacccaggtgttaccac 600
 DB 541 aagttggtgacccaggtgttaccac 600
 QY 601 ccagaaggtgtgtgttatacacaac 660
 DB 601 ccagaaggtgtgtgttatacacaac 660
 QY 661 tctgaattgtgtgacagacttgaagcacttccactgctgctgctgctgctgctgctgctgctgct 720
 DB 661 tctgaattgtgtgacagacttgaagcacttccactgctgctgctgctgctgctgctgctgctgct 720
 QY 721 gctagattggaagcacttgcacgtgttgaagcacttgcacgtgttgaagcacttgcacgtgttga 780
 DB 721 gctagattggaagcacttgcacgtgttgaagcacttgcacgtgttgaagcacttgcacgtgttga 780
 QY 781 atggacattgttccattgacactgttgcacacttgcacacttgcacacttgcacacttgcacacttgc 840
 DB 781 atggacattgttccattgacactgttgcacacttgcacacttgcacacttgcacacttgcacacttgc 840
 QY 841 tctctgactgttccactgac 900
 DB 841 tctctgactgttccactgac 900
 QY 901 aagttactcgttgcgt 960

DB 901 aagttactcgttgcgt 960
 QY 961 aacgaattgattgctagattgactcactctccagttccaaagacacacacttctactaacac 1020
 DB 961 aacgaattgattgctagattgactcactctccagttccaaagacacacacttctactaacac 1020
 QY 1021 actttgactcctaaaccagacttctccactgaaacgctacttctgacgtgacttctct 1080
 DB 1021 actttgactcctaaaccagacttctccactgaaacgctacttctgacgtgacttctct 1080
 QY 1081 cagcaaacacactatgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
 DB 1081 cagcaaacacactatgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
 QY 1141 ttgtctactacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1200
 DB 1141 ttgtctactacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1200
 QY 1201 gttccattcgtcgttagacttacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1260
 DB 1201 gttccattcgtcgttagacttacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1260
 QY 1261 ttggttagagtttgggttaacgacagattgttccattgcaacggttgggtgttgacaag 1320
 DB 1261 ttggttagagtttgggttaacgacagattgttccattgcaacggttgggtgttgacaag 1320
 QY 1321 ttggttagatgtaagagacgacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1380
 DB 1321 ttggttagatgtaagagacgacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1380
 QY 1381 aactgggaagaattgttctgcttaa 1404
 DB 1381 aactgggaagaattgttctgcttaa 1404

RESULT 8

AAZ31520

ID AAZ31520 standard; DNA; 1404 BP.

XX AC AAZ31520;

XX DT 06-JAN-2000 (first entry)

XX DE Consensus phytase-1-thermo(8)-Q50T-K91A coding sequence.

XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 consensus sequence; ds.

XX OS Synthetic.

XX PN WO9948380-A1.

XX PD 30-SEP-1999.

XX PF 22-MAR-1999; 99WO-DK00154.

XX PR 23-MAR-1998; 98DK-0000407.

XX PR 19-JUN-1998; 98DK-0000806.

XX PR 18-SEP-1998; 98DK-0001176.

XX PR 22-JAN-1999; 99DK-0000091.

XX PR 22-JAN-1999; 99DK-0000093.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Petersen S;

XX DR WPI: 1999-591030/50.

XX DR P-PSDB; AAY43169.

XX PT Preparing animal feed using a thermostable phytase

XX PS Example 3; Fig 9; 71pp; English.

OS Aspergillus fumigatus AFCC32722.
 OS Aspergillus fumigatus AFCC58128.
 OS Aspergillus fumigatus AFCC26906.
 OS Aspergillus fumigatus AFCC32239.
 OS Emericella nidulans.
 OS Talaromyces thermophilus AFCC20186.
 OS Myceliophthora thermophila.
 OS Synthetic.
 XX Key
 FH Location/Qualifiers
 FT 1..1404
 FT CDS
 FT /*tag= a
 FT /product= "Phytase-1-thermo[8]-Q50T-K91A"
 XX
 PN EP969089-AL.
 XX
 XX 05-JAN-2000.
 XX
 XX 23-JUN-1999; 99EP-0111949.
 XX
 XX 29-JUN-1998; 98EP-0111960.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 XX Brugger R, Lehmann M, Wyss M;
 PI
 XX WPI: 2000-099429/09.
 DR P-PSDB; AAY69568.
 DR
 XX New stabilized enzyme formulation, useful for feed compositions for
 PT monogastric animals.
 XX
 XX Example 5; Fig 19; 101pp; English.
 XX
 XX The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutamic and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents DNA encoding a mutant
 CC phytase-1 consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a
 CC temperature optimum and melting point 7 degrees Celsius higher than that
 CC of phytase-1 (AAY69558).
 XX
 XX Sequence 1404 BP; 329 A; 319 C; 304 G; 452 T; 0 other;
 SQ

Query Match 94.5%; Score 1327.2; DB 21; Length 1404;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1356; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 atggcggtgttcgtcgctactgtccattgccacattgttcggtccacatccggtacc 60
 |||||
 DB 1 atggcggtgttcgtcgctactgtccattgccacattgttcggtccacatccggtacc 60
 |||||
 QY 61 gcccttggtctcgttgtaactctactctgtgacactgttgacggtggtaccatgt 120
 |||||
 DB 61 gcccttggtctcgttgtaactctactctgtgacactgttgacggtggtaccatgt 120
 |||||

QY 121 ttccagaaatttctcaatttgggtggtacatactctccattctctcttcttggctgacgaa 180
 DB 121 ttccagaaatttctcaatttgggtggtacatactctccattctctcttcttggcagacgaa 180
 QY 181 tctgctatttctccagacgttccaaagggtttagaggttacttcttcttgaagtttctct 240
 DB 181 tctgctatttctccagacgttccagacgtttagaggttacttcttcttgaagtttctct 240
 QY 241 agacaggtgctagatatacccaacttcttctgctgcttaagcgtactctctcttgaatgaa 300
 DB 241 agacaggtgctagatatacccaacttcttctgctgcttaagcgtactctcttgaatgaa 300
 QY 301 gctattcaaaagaacgctactgctttcaagggttaagtacgcttcttcttgaagacttacaac 360
 DB 301 gctattcaaaagaacgctactgctttcaagggttaagtacgcttcttcttgaagacttacaac 360
 QY 361 tacactttgggtgctgacgacttgaacttccattcgggtgaacacaaatggttaactctggt 420
 DB 361 tacactttgggtgctgacgacttgaacttccattcgggtgaacacaaatggttaactctggt 420
 QY 421 attaagttctacagaagatacaaggctttgctgctagaaagattgttccattcattagagct 480
 DB 421 attaagttctacagaagatacaaggctttgctgctagaaagattgttccattcattagagct 480
 QY 481 tctggttctgacagagtattgtcttctgctgctgctgctgctgctgctgctgctgctgct 540
 DB 481 tctggttctgacagagtattgtcttctgctgctgctgctgctgctgctgctgctgctgct 540
 QY 541 aagttgctgacccagggtgcttaacccacacacacacacacacacacacacacacacacacac 600
 DB 541 aagttgctgacccagggttctcaacacacacacacacacacacacacacacacacacacac 600
 QY 601 ccacaaaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 660
 DB 601 ccacaaaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 660
 QY 661 tctgaattggtgacgacgttgaagcgttaacttcaacttcaacttcaacttcaacttcaacttca 720
 DB 661 tctgaattggtgacgacgttgaagcgttgaagcgttaacttcaacttcaacttcaacttcaacttca 720
 QY 721 gctagattggaagcgtgactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 780
 DB 721 gctagattggaagcgtgactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 780
 QY 781 atgacatgttccattcgcacactgttgcctgctgctgctgctgctgctgctgctgctgctgct 840
 DB 781 atgacatgttccattcgcacactgttgcctgctgctgctgctgctgctgctgctgctgctgct 840
 QY 841 tctgtgactgttccattcgcacacgaaatggaatggaatggaatggaatggaatggaatggaatg 900
 DB 841 tctgtgactgttccattcgcacacgaaatggaatggaatggaatggaatggaatggaatggaatg 900
 QY 901 aagttactacggttgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 960
 DB 901 aagttactacggttgcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 960
 QY 961 aacgaattgattgctgattgactgacttccacttccacttccacttccacttccacttccacttcc 1020
 DB 961 aacgaattgattgctgattgactgacttccacttccacttccacttccacttccacttccacttcc 1020
 QY 1021 accttggacttaacccagctacttccattgaaacgtacttcttctgctgctgctgctgctgctgct 1080
 DB 1021 accttggacttaacccagctacttccattgaaacgtacttcttctgctgctgctgctgctgctgct 1080
 QY 1081 caacgaacacactatggttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
 DB 1081 caacgaacacactatggttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
 QY 1141 ttgtctactacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1200
 DB 1141 ttgtctactacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1200

QY 841 ttctgtgacttgcactcaacgaatggattcaatacagactacttgcacttcttgggt 900
 Db 841 ttctgtgacttgcactcaacgaatggattcaatacagactacttgcacttcttgggt 900
 QY 901 aagctacaggttacggtctgttaaccattgggtccagctcaagggttgggttcgtt 960
 Db 901 aagctacaggttacggtctgttaaccattgggtccagctcaagggttgggttcgtt 960
 QY 961 aacgaattgattgtagtactcactctccagttccaggttcaagaccacacttactaaccac 1020
 Db 961 aacgaattgattgtagtactcactctccagttccaggttcaagaccacacttactaaccac 1020
 QY 1021 acttggacttaaccagactacttccattgaacgtacttctgacgtgacttctct 1080
 Db 1021 acttggacttaaccagactacttccattgaacgtacttctgacgtgacttctct 1080
 QY 1081 cagcaaacactatggttcttcttctcgtcttgggtttgtacacaggtactaagcca 1140
 Db 1081 cagcaaacactatggttcttcttctcgtcttgggtttgtacacaggtactaagcca 1140
 QY 1141 ttgtctactacttctgttgaaactattgaagaactgacggttactctgcttcttggact 1200
 Db 1141 ttgtctactacttctgttgaaactattgaagaactgacggttactctgcttcttggact 1200
 QY 1201 gtccattcgtctgtagacttacttgaatgatgcaatgtgaagtgaaagaaacca 1260
 Db 1201 gtccattcgtctgtagacttacttgaatgatgcaatgtgaagtgaaagaaacca 1260
 QY 1261 ttggttagagtttggtaacgacagagttgttccattgacggttgggttgacaaag 1320
 Db 1261 ttggttagagtttggtaacgacagagttgttccattgacggttgggttgacaaag 1320
 QY 1321 ttggttagatgtagagagacgactctgtgaaggtttgtcttctgtagatctgggtt 1380
 Db 1321 ttggttagatgtagagagacgactctgtgaaggtttgtcttctgtagatctgggtt 1380
 QY 1381 aactgggaagaatttccgcttaa 1404
 Db 1381 aactgggtgtagtcttccgcttaa 1404

RESULT 13

ID AA227423
 ID AA227423 standard; cDNA; 1426 BP.

AC AA227423;
 XX

DT 07-DEC-1999 (first entry)
 XX

DE Ascomycete consensus phytase coding sequence.
 XX

KW Phytase; variant; enzyme; phosphorus liberation; phytase substrate;
 KW Phytase level reduction; animal manure; food preparation;
 KW soy processing; inositol manufacture; ss.

OS Synthetic.
 XX

PN WO9949022-A1.
 XX

PD 30-SEP-1999.
 XX

PF 22-MAR-1999; 99WO-DK00153.
 XX

PR 23-MAR-1998; 98DK-0000407.
 PR 19-JUN-1998; 98DK-0000806.

PR 18-SEP-1998; 98DK-0001176.
 PR 22-JAN-1999; 99DK-0000091.

XX (NOVO) NOVO-NORDISK AS.
 PA

XX Svendsen A;
 PI

XX WPI; 1999-580444/49.
 XX

DB

P-PSDB; AAY39906.

XX New variant phytase enzymes, used for liberating phosphorus from a
 PT phytase substrate, for reducing phytate levels in animal manure and in
 PT feed and food preparations -
 XX Disclosure; Fig 9f-g; 141pp; English.

XX This sequence encodes the consensus Ascomycete phytase sequence.

CC The invention relates to variant phytase enzymes with specific amino acid
 CC substitutions for improved properties. The phytase variants can be used
 CC for liberating phosphorus from a phytase substrate. They can be used for
 CC reducing phytate levels in animal manure. They can be used in feed or
 CC food preparations. The phytase DNA can also be used to produce transgenic
 CC plants which can be used in feeds or foods. The phytase variants can also
 CC be used in soy processing and in the manufacture of inositol or
 CC derivatives. The phytase variants can have altered activities such as pH
 CC stability, temperature stability, pH profile, temperature profile,
 CC specific activity (in particular in relation to pH and temperature),
 CC substrate specificity, substrate cleavage pattern, substrate binding,
 CC position specificity, the velocity and level of release of phosphate
 CC from corn, reaction rate, phytase degradation rate and end level of
 CC released phosphate reached.

XX

SQ Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;

Query Match 93.8%; Score 1317.6; DB 20; Length 1426;

Best Local Similarity 96.2%; Pred. No. 0;

Matches 1350; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 atggcggtgtctgctgacttccattgcacattgcacattgttgcgttccacatccggtacc 60

Db 12 atggcggtgtctgctgacttccattgcacattgcacattgttgcgttccacatccggtacc 71

QY 61 gccttgggtctcgttgtaactctcactctgtgacactgttgacactgttgacactgttgcgttacc 120

Db 72 gccttgggtctcgttgtaactctcactctgtgacactgttgacactgttgacactgttgcgttacc 131

QY 121 tcccaagaatttctcacttgggtgggtacatactctccattctcttcttcttgcgtgacgaa 180

Db 132 tcccaagaatttctcacttgggtgggtacatactctccattctcttcttcttgcgtgacgaa 191

QY 181 tctgtatttctccagacgttccaaagggtttagagttacttcttcttgcgttaccagtgtgtct 240

Db 192 tctgtatttctccagacgttccaaagggtttagagttacttcttcttgcgttaccagtgtgtct 251

QY 241 agacacggtgtgtagataccacacttcttctgcttcttaaggcgtactcttcttgcgttacc 300

Db 252 agacacggtgtgtagataccacacttcttctgcttcttaaggcgtactcttcttgcgttacc 311

QY 301 gctattcaaaagacgtactcttcttcaagggttaagtagcgttcttcttgaagacttaccac 360

Db 312 gctattcaaaagacgtactcttcttcaagggttaagtagcgttcttcttgaagacttaccac 371

QY 361 tacacttgggtgctgacgacttgcctcattcgttgcgttcttaaggcgtactcttcttgcgttacc 420

Db 372 tacacttgggtgctgacgacttgcctcattcgttgcgttcttaaggcgtactcttcttgcgttacc 431

QY 421 attagttctcagaagatatacaggcttcttgcgttagaagattgttccattcattagact 480

Db 432 attagttctcagaagatatacaggcttcttgcgttagaagattgttccattcattagact 491

QY 481 tctgttctcagagagttattgtcttctcaggttaaggttcaattgaaggttcccaatctgct 540

Db 492 tctgttctcagagagttattgtcttctcaggttaaggttcaattgaaggttcccaatctgct 551

QY 541 aagttgggtgacctgctgcttaaccac 600

Db 552 aagttgggtgacctgctgcttaaccac 611

QY 601 ccagaaggtgctggttatacaacaacacacttggaccacggttctgttactgcttccgaaagaa 660

Db 612 ccagaaagatccggttacacaacactttggaccacggtactgtactgtcttgcgaagac 671
 Qy 661 tctgaattgggtgacagcgttgaagtaacttaacttcaactgctgttttgcgtccccaattaga 720
 Db 672 tctgaattgggtgacagcgttgaagtaacttaacttcaactgctgttttgcgtccccaattaga 731
 Qy 721 gctagattggaagctcaacttcccgaggttgaactgaactgacgaagcgttgaacttg 780
 Db 732 gctagattggaagctgacttgcaggttgaacttgaactgacgaagcgttgaacttg 791
 Qy 781 atggacatgtccattcgacactgttctagaacttctgacgttcaacttctctcca 840
 Db 792 atggacatgtccattcgacactgttctagaacttctgacgttcaacttctctcca 851
 Qy 841 tctgtgactgttccactcagcagaatggattcaactgaactgaacttgaacttcttgggt 900
 Db 852 tctgtgactgttccactcagcagaatggattcaactgaactgaacttgaacttcttgggt 911
 Qy 901 aagtaactcaggtgacttgaacttgaacttgaacttgaacttgaacttgaacttgaact 960
 Db 912 aagtaactcaggtgacttgaacttgaacttgaacttgaacttgaacttgaacttgaact 971
 Qy 961 aacgaattgattgctagattgacttcaacttccagttcaagacacacttctactaaccac 1020
 Db 972 aacgaattgattgctagattgacttcaacttccagttcaagacacacttctactaaccac 1031
 Qy 1021 acttggacttaaccagacttcttccactgaactgaactgaactgaactgaactgaacttct 1080
 Db 1032 acttggacttaaccagacttcttccactgaactgaactgaactgaactgaactgaacttct 1091
 Qy 1081 cagcaaacactatgatttcttcttcttcttcttcttcttcttcttcttcttcttctt 1140
 Db 1092 cagcaaacactatgatttcttcttcttcttcttcttcttcttcttcttcttcttctt 1151
 Qy 1141 tctgtactacttctgttgaattatgttaagaactgacggttacttcttcttcttcttctt 1200
 Db 1152 tctgtactacttctgttgaattatgttaagaactgacggttacttcttcttcttcttctt 1211
 Qy 1201 gttccattcgtctagacttcaacttgaactgaactgaactgaactgaactgaactgaact 1260
 Db 1212 gttccattcgtctagacttcaacttgaactgaactgaactgaactgaactgaactgaact 1271
 Qy 1261 tctgttagagtttgggttaacagacagagtttcttccattgacggttgggttgggtgacaa 1320
 Db 1272 tctgttagagtttgggttaacagacagagtttcttccattgacggttgggttgggtgacaa 1331
 Qy 1321 tctgttagagtttgggttaacagacagacttcttgaaggttcttcttcttcttcttctt 1380
 Db 1332 tctgttagagtttgggttaacagacagacttcttgaaggttcttcttcttcttcttctt 1391
 Qy 1381 aactgggaagaatgttctcttaa 1404
 Db 1392 aactgggaagaatgttctcttaa 1415

RESULT 14

AA23022

ID AA23022 standard; DNA; 1426 BP.

XX AC AAX23022;

XX AC AAX23022;

DT 11-JUN-1999 (first entry)

XX Fungal phytase gene consensus DNA.

XX Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; mutin; feed; pharmaceutical; ds.
 OS Fungi.

XX Key

XX Location/Qualifiers
 FT 12..1415
 FT /tag= a

/product= "phytase"

FT XX

PN EP897985-A2.

XX XX

PD 24-FEB-1999.

XX XX

PF 15-JUL-1998; 98EP-0113176.

XX XX

PR 24-JUL-1997; 97EP-0112688.

XX XX

PA (HOFF) HOFFMANN LA ROCHE AG F.

XX XX

PI Lehmann M;

XX XX

DR WPI; 1999-134647/12.

XX XX

DR P-PSDB; AAW33380.

XX XX

PT Preparation of a consensus protein, especially a phytase - using

XX programs to compare evolutionary similarity of sequences

PS Claim 8; Fig 2; 30pp; English.

XX

CC This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses
 CC phytase to valuable feed additives, with a fully defined amino acid
 CC sequence given in the specification, or variant or mutin. The method is
 CC useful for improving protein properties by altering their sequence. The
 CC consensus protein and mutin are useful in food, feed or pharmaceutical
 CC compositions. This sequence encodes the consensus phytase protein used
 CC in the method of the invention.

XX Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;
 SQ

Query Match

93.8%; Score 1317.6; DB 20; Length 1426;

Best Local Similarity 96.2%; Pred. No. 0;

Matches 1350; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 1 atggcgctgttcgctgactgctccattgcccacttcttgcgttccacatccggtacc 60

Db 12 atggcgctgttcgctgactgctccattgcccacttcttgcgttccacatccggtacc 71

Qy 61 gccttgggtctctggttaacttcaacttcttgcacttcttgcacttcttgcacttcttgcactt 120

Db 72 gccttgggtctctggttaacttcaacttcttgcacttcttgcacttcttgcacttcttgcactt 131

Qy 121 tcccaaaaaatttctcaacttcttgcgttgcacttcttgcacttcttgcacttcttgcactt 180

Db 132 tcccaaaaaatttctcaacttcttgcgttgcacttcttgcacttcttgcacttcttgcactt 191

Qy 181 tctgtatttctccagacgttcccaagggttgcacttcttgcacttcttgcacttcttgcactt 240

Db 192 tctgtatttctccagacgttcccaagggttgcacttcttgcacttcttgcacttcttgcactt 251

Qy 241 agacacggtgctagatatacccaacttcttgcacttcttgcacttcttgcacttcttgcactt 300

Db 252 agacacggtgctagatatacccaacttcttgcacttcttgcacttcttgcacttcttgcactt 311

Qy 301 gctattcaaaaagacgttacttcttcaagggttaagtaacgttcttcttgcacttcttgcactt 360

Db 312 gctattcaaaaagacgttacttcttcaagggttaagtaacgttcttcttgcacttcttgcactt 371

Qy 361 tacacttgggtgctgacgacttgaacttcttgcacttcttgcacttcttgcacttcttgcactt 420

Db 372 tacacttgggtgctgacgacttgaacttcttgcacttcttgcacttcttgcacttcttgcactt 431

Qy 421 attaatgtctcagaagatacaaggcttcttgcacttcttgcacttcttgcacttcttgcactt 480

Db 432 attaatgtctcagaagatacaaggcttcttgcacttcttgcacttcttgcacttcttgcactt 491

Qy 481 tctgttctgcagaggttattgcttcttgcacttcttgcacttcttgcacttcttgcactt 540

Db 492 tctgtgtgacagagttattgtcttctgtgaaagttcaattgaagggtttccaattgct 551
 QY 541 aagttgggtgacccaggtgttaacccacacaaagcttctccagttatcaacgttatatt 600
 Db 552 aagttgggtgacccaggtgttaacccacacaaagcttctccagttatgaagttatt 611
 QY 601 ccagaaggtgctggttaacacacacactttgacacacaggtttgttactgctttcgaagaa 660
 Db 612 ccagaaggtgctggttaacacacacactttgacacacaggtttgttactgctttcgaagac 671
 QY 661 tctgaattgggtgacagcgttgaaagcttaactcactgctgttttctgcctccacaaattaga 720
 Db 672 tctgaattgggtgacagcgttgaaagcttaactcactgctgttttctgcctccacaaattaga 731
 QY 721 gctagattggaagctgacttgcaggtgttaacttgaactgacgaagcgtttgtaacttg 780
 Db 732 gctagattggaagctgacttgcaggtgttaacttgaactgacgaagcgtttgtaacttg 791
 QY 781 atggacatgtgtccattgcacactgttctgaacttctgaacttctgaacttctgaacttctga 840
 Db 792 atggacatgtgtccattgcacactgttctgaacttctgaacttctgaacttctgaacttctga 851
 QY 841 tctgtgacttgttcaactcagcagaaatgattcaatagcactacttgcaacttctgggt 900
 Db 852 tctgtgacttgttcaactcagcagaaatgattcaatagcactacttgcaacttctgggt 911
 QY 901 aagttactacgttgcaggtgttgcagcattggttgcagcattggttgcagcattggttgcag 960
 Db 912 aagttactacgttgcaggtgttgcagcattggttgcagcattggttgcagcattggttgcag 971
 QY 961 aacgaattgattgttagattgactcacttccagttccagttccagttccagttccagttccag 1020
 Db 972 aacgaattgattgttagattgactcacttccagttccagttccagttccagttccagttccag 1031
 QY 1021 actttggacttaaccagctacttccagttccagttccagttccagttccagttccagttccag 1080
 Db 1032 actttggacttaaccagctacttccagttccagttccagttccagttccagttccagttccag 1091
 QY 1081 cagcaaacactatgtttctatttcttctgttgggtttgtacacaggttacttaagcca 1140
 Db 1092 cagcaaacactatgtttctatttcttctgttgggtttgtacacaggttacttaagcca 1151
 QY 1141 ttgtctactacttctgttgaatctattgaagaaactgacggttacttctgttcttggact 1200
 Db 1152 ttgtctactacttctgttgaatctattgaagaaactgacggttacttctgttcttggact 1211
 QY 1201 gttccattctgctgctagagcttaacttgaatgatgcaatgtgaagctgaaaggaacca 1260
 Db 1212 gttccattctgctgctagagcttaacttgaatgatgcaatgtgaagctgaaaggaacca 1271
 QY 1261 ttggttagagtttgggttaacagcagaggttgcattgacaggttgggttgggttgcagag 1320
 Db 1272 ttggttagagtttgggttaacagcagaggttgcattgacaggttgggttgggttgcagag 1331
 QY 1321 ttggttagagtttgggttaacagcagaggttgcattgacaggttgggttgggttgcagag 1380
 Db 1332 ttggttagagtttgggttaacagcagaggttgcattgacaggttgggttgggttgcagag 1391
 QY 1381 aactgggaagaattgttctgcttaa 1404
 Db 1392 aactgggctgaattgttctgcttaa 1415

RESULT 15

ID AAA73231 standard; DNA; 1426 BP.
 XX

AC AAA73231;
 XX

DT 05-DEC-2000 (first entry)
 XX

DE Consensus phytase polynucleotide sequence SEQ ID NO:15.
 XX

KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure; ds.
 XX

OS Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX P-PSDB; AAB20515.

Novel phytases with improved properties such as temperature stability,
 pH stability and substrate specificity, for use in pharmaceuticals and
 compound foods and feeds -

Example 9; Fig 2a-c; 240pp; English.

The present invention describes improved phytases, preferably with
 increased thermostability, and methods for producing them. The methods
 can be used for producing phytases with improved properties e.g.
 temperature stability, pH stability, pH profile, temperature profile,
 specific activity, substrate specificity, substrate cleavage pattern,
 substrate binding, position specificity, the velocity and level of
 release of phosphate from corn, reaction rate, phytate degradation rate,
 and end level of released phosphate. The phytases can be used to produce
 pharmaceutical compositions or compound food or feeds. The feed can be
 used to reduce levels of phytate in animal manure, by converting it
 into lower inositol phosphates and/or inositol and inorganic phosphate.
 The present sequence encodes a phytase sequence from the present
 invention.

Sequence 1426 BP; 338 A; 311 C; 308 G; 469 T; 0 other;

Query Match 93.8%; Score 1317.6; DB 21; Length 1426;

Best Local Similarity 96.2%; Pred. No. 0;

Matches 1350; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 atggcggtgttgcgtgctactgtccacttgcacacttgcgttccacacccgggtacc 60

Db 12 atggcggtgttgcgtgctactgtccacttgcacacttgcgttccacacccgggtacc 71

QY 61 gcttgggttccctggtgtaactctcacttctgtgacactgttgacggttgacaaagt 120

Db 72 gcttgggttccctggtgtaactctcacttctgtgacactgttgacggttgacaaagt 131

QY 121 ttcccagaaattctcacttgcgttgggtgacatactctcacttctcttcttggcgtaacaa 180

Db 132 ttcccagaaattctcacttgcgttgggtgacatactctcacttctcttcttggcgtaacaa 191

QY 181 tctgtatttctccagacgttccaaagggtgtgtagagttacttctgtaagtttctgt 240

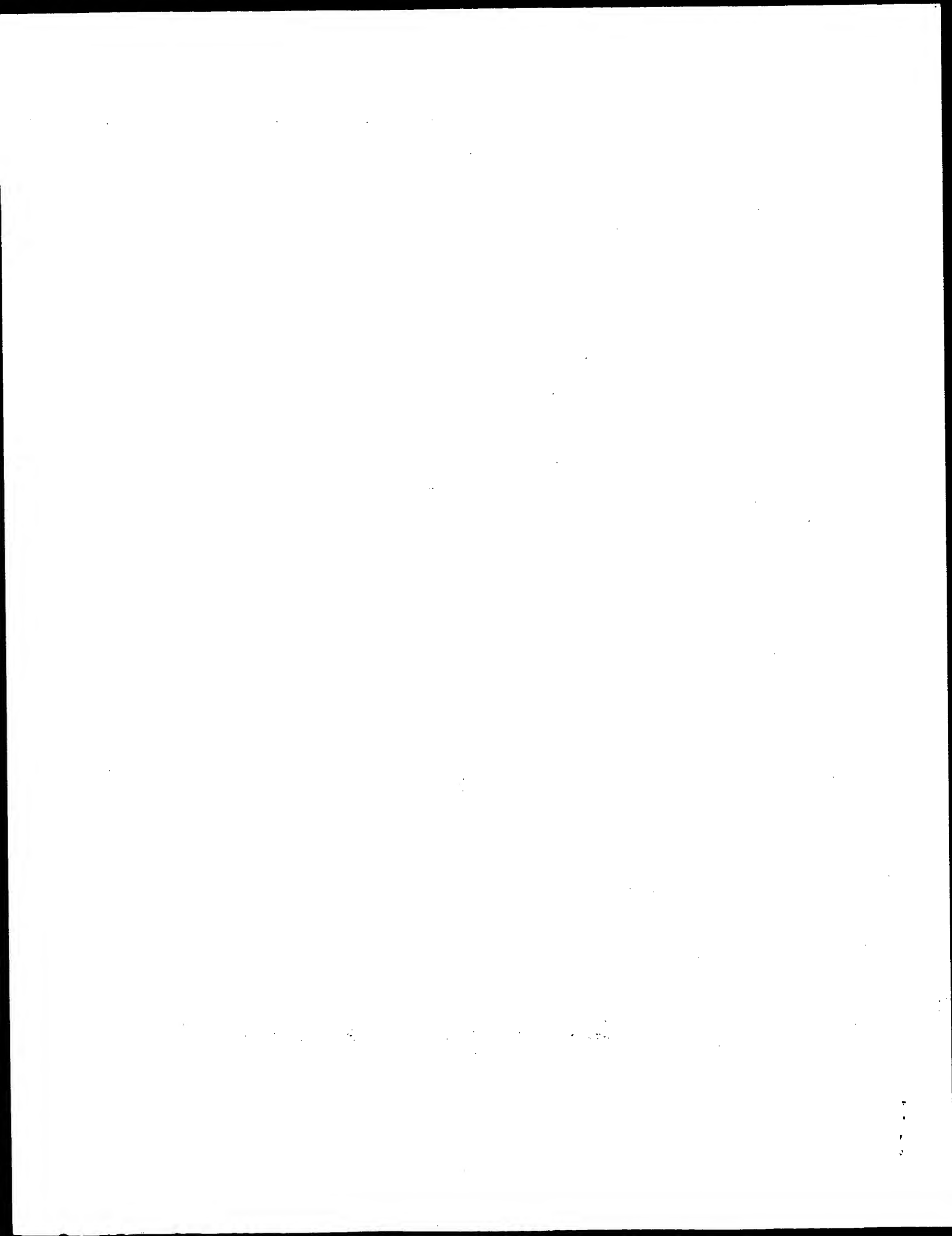
Db 192 tctgtatttctccagacgttccaaagggtgtgtagagttacttctgtaagtttctgt 251

QY 241 agacacggtgtctagataccacacttctctgctcactaagcggtactctgttattgaa 300

Db 252 agacacggtgtctagataccacacttctctgctcactaagcggtactctgttattgaa 311

QY 301 gctattccaaagacgctactgcttccaaagggtgaagtacgcttcttgaagacttaaac 360

Db 312 gctattcaaaagacgcgtactgcttctcaagggttaagtaacgcttcttctgaagacttaacaac 371
QY 361 tacactttgggtgctgacgacttgactccattcggtgaacacaaatgggttaactctggt 420
Db 372 tacactttgggtgctgacgacttgactccattcggtgaacacaaatgggttaactctggt 431
QY 421 attaaagtctacagaagaatacaaggcttggctagaagaattgtccattcatttagagct 480
Db 432 attaaagtctacagaagaatacaaggcttggctagaagaattgtccattcatttagagct 491
QY 481 tctggttctgacagagttattgcttctgctgaagaattcattgaaggtttccaaatctgct 540
Db 492 tctggttctgacagagttattgcttctgctgaagaattcattgaaggtttccaaatctgct 551
QY 541 aagttggtgacccaggtgcttaaccacacccaagcttctccagttataaacggttattatt 600
Db 552 aagttggtgacccaggttctcaaccacccaagcttctccagttattgaaggttattatt 611
QY 601 ccagaaggtgctgttatacaacaacacttggaccacaggtttgtgactgctttcgaagaa 660
Db 612 ccagaaggtatccggtatacaacaacacttggaccacaggttctgactgctttcgaagac 671
QY 661 tctgaattgggtgacgagcttgaagtaacttcaactcactgcttttctgctccaccaattaga 720
Db 672 tctgaattgggtgacgagcttgaagtaacttcaactcactgcttttctgctccagctattaga 731
QY 721 gctagattggaagctcaacttccaggtgttaacttgactgacgaagacggttggtaacttg 780
Db 732 gctagattggaagctgacttgcaggtgttacttgaactgacgaagacggttggtaacttg 791
QY 781 atggacatgttccattcgacacttctgctagaacttctgactgactcaacttctctcca 840
Db 792 atggacatgttccattcgacacttctgctagaacttctgactgactcaacttctctcca 851
QY 841 tctgtgacttgttcaactcaacgaacgaattggattcaatacgaactacttgcacttcttgggt 900
Db 852 tctgtgcttgttcaactcaacgaacgaattggagacaactacgactacttgcacttcttgggt 911
QY 901 aagttactacggttacggtgctgttaaccattgggtccagctcaaggtgttgggttctggt 960
Db 912 aagttactacggttacggtgctgtgttaaccattgggtccagctcaaggtgttgggttctggt 971
QY 961 aacgaattgattgctagattgactcacttccagttccagttcaagaccacacttctactaaccc 1020
Db 972 aacgaattgattgctagattgactcacttccagttccagttcaagaccacacttctactaaccc 1031
QY 1021 actttggacttaaccacgactacttccattccattgaacgctacttttgaacgctgacttctct 1080
Db 1032 actttggacttaaccacgactacttccattccattgaacgctacttttgaacgctgacttctct 1091
QY 1081 cagcaacacactatggtttcttcttcttctgcttgggtttgttacaacggtacttaagcca 1140
Db 1092 cagcaacactctatgattcttcttcttcttctgcttgggtttgttacaacggtacttctcca 1151
QY 1141 ttgtctactacttctgttgaattctattgaagaaactgacggttacttctgcttcttggact 1200
Db 1152 ttgtctactacttctgttgaattctattgaagaaactgacggttacttctgcttcttggact 1211
QY 1201 gttccattcgctgctagagcttcaacttgaatgtagtgaatgtagcgtgaaaggaacca 1260
Db 1212 gttccattcggtgctagagcttcaacttgaatgtagtgaatgtagcgtgaaaggaacca 1271
QY 1261 ttggttagagtttctggttaacgacagagttgtccattgacggttgggtgttgacaag 1320
Db 1272 ttggttagagtttctggttaacgacagagttgtccattgacggttgggtgttgacaag 1331
QY 1321 ttggttagatgtaagagagacgacttctggtgaaggttcttcttctgctagatctggtggt 1380
Db 1332 ttggttagatgtaagagagacgacttctggtgaaggttcttcttctgctagatctggtggt 1391
QY 1381 aactgggaagaatgtttctgcttaa 1404
Db 1392 aactgggctgaatgtttctgcttaa 1415



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 26, 2001, 15:20:18 ; Search time 87.5 Seconds
(without alignments)
3037.632 Million cell updates/sec

Title: US-09-488-265-30

Perfect score: 1404
Sequence: 1 atgggggtgttcgtgct.....gggaagaatgttcgttaa 1404

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1317.6	93.8	1426	3	US-09-121-425-3
2	614.2	43.7	1404	1	Sequence 3, Appli
3	614.2	43.7	1404	1	Sequence 33, Appl
4	614.2	43.7	1404	1	Sequence 19, Appl
5	614.2	43.7	1404	2	Sequence 1, Appli
6	610.2	43.5	2363	1	Sequence 33, Appl
7	610.2	43.5	2363	2	Sequence 7, Appli
8	610.2	43.5	2379	2	Sequence 7, Appli
9	610.2	43.5	6756	1	Sequence 1, Appli
10	610.2	43.5	6756	2	Sequence 31, Appl
11	591.6	42.1	1515	3	Sequence 31, Appl
12	590.6	42.1	1332	3	Sequence 5, Appli
13	346	24.6	2280	4	Sequence 4, Appli
14	346	24.6	2200	4	Sequence 1, Appli
15	131.4	9.4	1320	3	Sequence 1, Appli
16	131.4	9.4	1320	3	Sequence 1, Appli
17	131.4	9.4	1593	3	Sequence 1, Appli
18	112.2	8.0	1501	3	Sequence 23, Appl
19	100	7.1	1522	3	Sequence 21, Appl
20	90.8	6.5	1642	3	Sequence 25, Appl
21	87.6	6.2	1536	3	Sequence 27, Appl
22	48.4	3.4	7218	1	Sequence 29, Appl
23	46.8	3.3	99	1	Sequence 14, Appl
24	46.8	3.3	99	2	Sequence 12, Appl
25	42.4	3.0	1631	2	Sequence 12, Appl
26	42.4	3.0	1631	2	Sequence 5, Appli
27	42.4	3.0	1631	3	Sequence 5, Appli

28 35.2 2.5 1867 1 US-07-772-087-1 Sequence 1, Appli
29 34 2.4 2126 4 US-08-235-836C-75 Sequence 75, Appl
30 33 2.4 14507 3 US-08-785-150-1 Sequence 1, Appli
31 32.8 2.3 2165 1 US-08-222-289-1 Sequence 1, Appli
32 32.6 2.3 585 3 US-08-882-501-28 Sequence 28, Appl
33 32.2 2.3 1360 3 US-08-961-083-37 Sequence 37, Appl
34 32 2.3 8367 2 US-08-583-562B-7 Sequence 7, Appli
35 32 2.3 8367 2 US-08-779-113-7 Sequence 31, Appl
36 31.8 2.3 276 3 US-08-993-359-31 Sequence 31, Appl
37 31.2 2.2 720 3 US-09-479-309-4 Sequence 4, Appli
38 31.2 2.2 1404 3 US-08-961-083-9 Sequence 9, Appli
39 31.2 2.2 2601 3 US-09-039-773A-2 Sequence 2, Appli
40 31 2.2 3084 1 US-08-551-437-1 Sequence 1, Appli
41 31 2.2 3084 3 US-09-004-225-1 Sequence 1, Appli
42 31 2.2 3084 4 US-09-084-346-1 Sequence 1, Appli
43 30.8 2.2 66 1 US-08-151-574-11 Sequence 11, Appl
44 30.8 2.2 66 2 US-08-419-448-11 Sequence 11, Appl
45 30.8 2.2 3147 4 US-09-066-046-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-121-425-3
; Sequence 3, Application US/09121425
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121,425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1426
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-3

Query Match 93.8%; Score 1317.6; DB 3; Length 1426;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1350; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 atgggggtgttcgtgctgtaactgtccattgcaactgttgcgttccacatcccggtacc 60
Db 12 atgggggtgttcgtgctgtaactgtccattgcaactgttgcgttccacatcccggtacc 71
QY 61 gcttgggttcctcggtgtaactgtccattgcaactgttgcgttccacatcccggtacc 120
Db 72 gcttgggttcctcggtgtaactgtccattgcaactgttgcgttccacatcccggtacc 131
QY 121 ttccagaaatttcacattgtgggtacatactctccattctctcttcttggctgacgaa 180
Db 132 ttccagaaatttcacattgtgggtacatactctccattctctcttcttgggtgacgaa 191
QY 181 tctgtatttccagagcttccaaaggttgcagacttctctgtaagggtactctgcttgcattgct 240
Db 192 tctgtatttccagagcttccaaaggttgcagacttctctgtaagggtactctgcttgcattgct 251
QY 241 agacacggtgctagataccacattctctgctgtaagggtactctgcttgcattgctgaa 300
Db 252 agacacggtgctagataccacattctctgctgtaagggtactctgcttgcattgctgaa 311
QY 301 gctattcaaaagacgctactgttccaaaggtgtaagggtactcttcttgaagacttcaaac 360

[illegible]

NAME/KEY: Signal Sequence
 LOCATION: 1...72
 OTHER INFORMATION:
 US-08-693-709-1

Query Match 43.7%; Score 614.2; DB 1; Length 1404;
 Best Local Similarity 64.9%; Pred. No. 4.6e-173;
 Matches 910; Conservative 0; Mismatches 493; Indels 0; Gaps 0;

```

Qy 1 atggcggtgttcgtgctgactgtccattgccacccttggttcggttcacatcccggtacc 60
Db 1 atggcggtgttcgtgctgactgtccattgccacccttggttcggttcacatcccggtacc 60
Qy 61 gcttgggtgctcgtggtgaaactctcaactcttgtagacactgttgacggtggtttacaatgt 120
Db 61 gcttgggtgctcgtggtgaaactctcaactcttgtagacactgttgacggtggtttacaatgt 120
Qy 121 ttccagaaattcttbaacttggtgggtacatactctccattctctcttcttggcgtgacgaa 180
Db 121 ttccagaaattcttbaacttggtgggtacatactctccattctctcttcttggcgtgacgaa 180
Qy 181 tctgctattctccagactgtccaaagggttgtagagtaactttcgttcaagtttgtct 240
Db 181 tctgctattctccagactgtccaaagggttgtagagtaactttcgttcaagtttgtct 240
Qy 241 agacacggtgctagatatacccaactcttctcgtcgtctaaggcgtactcgtgcttgattgaa 300
Db 241 agacacggtgctagatatacccaactcttctcgtcgtctaaggcgtactcgtgcttgattgaa 300
Qy 301 gctattcaaaaagcgtactgtcttcaagggtgaagtacgcttcttctgaagacttacac 360
Db 301 gctattcaaaaagcgtactgtcttcaagggtgaagtacgcttcttctgaagacttacac 360
Qy 361 taactttgggtgctgacgactgactccattcgggtgaacaaatggttaactctggt 420
Db 361 taactttgggtgctgacgactgactccattcgggtgaacaaatggttaactctggt 420
Qy 421 attaagttctacagaagatacaaggcttcttgctagaagaattgttccattcattagact 480
Db 421 attaagttctacagaagatacaaggcttcttgctagaagaattgttccattcattagact 480
Qy 481 tctggtttgacagagtattgtcttctgctgtaaaagtctcattgaaggtttccaaatcgtct 540
Db 481 tctggtttgacagagtattgtcttctgctgtaaaagtctcattgaaggtttccaaatcgtct 540
Qy 541 aagttggtgacccaggtgttaaccacacacacacacacacacacacacacacacacacac 600
Db 541 aagttggtgacccaggtgttaaccacacacacacacacacacacacacacacacacacac 600
Qy 601 ccagaaggtgctggtttacacacacacacacacacacacacacacacacacacacacacac 660
Db 601 ccagaaggtgctggtttacacacacacacacacacacacacacacacacacacacacacac 660
Qy 661 tctgaattgggtgacgactgaaacttaactcactgctgttttgcgtccaccacaaataga 720
Db 661 tctgaattgggtgacgactgaaacttaactcactgctgttttgcgtccaccacaaataga 720
Qy 721 gctagattggaagtcactgccaaggtgttaacttgactgacgacgaagcgtgttaacttg 780
Db 721 gctagattggaagtcactgccaaggtgttaacttgactgacgacgaagcgtgttaacttg 780
Qy 781 atggacatgtgtccattgcacactgtgtgtagaactctgacgtactcaattgtctcca 840
Db 781 atggacatgtgtccattgcacactgtgtgtagaactctgacgtactcaattgtctcca 840
Qy 841 tctgtgactgttctcactcagcagaatgatttcaatcacgactacttgcaactcttgggt 900
Db 841 tctgtgactgttctcactcagcagaatgatttcaatcacgactacttgcaactcttgggt 900
Qy 901 aagtaactaggtgtacggtgtgttaaccacacacacacacacacacacacacacacacac 960
Db 901 aagtaactaggtgtacggtgtgttaaccacacacacacacacacacacacacacacacac 960
Qy 961 aacgaattgattgctagattgactcaactctccagttcaagaccacacacacttctactaacac 1020
Db 961 aacgaattgattgctagattgactcaactctccagttcaagaccacacacacttctactaacac 1020
Qy 1021 actttggacttaaccagcacttctccattgaacgtcaactttgtacgctgacttctct 1080
Db 1021 actttggacttaaccagcacttctccattgaacgtcaactttgtacgctgacttctct 1080
Qy 1081 cagcaaacacatattggtttctattttcttctcgttcttgggtttgtacaaaggttactaagcca 1140
Db 1081 cagcaaacacatattggtttctattttcttctcgttcttgggtttgtacaaaggttactaagcca 1140
Qy 1141 tigtctactactctcgttgaaactattgaagaacacgtgacgttactcgttcttcttggact 1200
Db 1141 tigtctactactctcgttgaaactattgaagaacacgtgacgttactcgttcttcttggact 1200
Qy 1201 gtccattcgtcgttagagcttacgttgaatgatgcaatgtgaagctgaagaagaaacca 1260
Db 1201 gtccattcgtcgttagagcttacgttgaatgatgcaatgtgaagctgaagaagaaacca 1260
Qy 1261 ttggttagattttggtttaacgacagagtttccattcgcacggttggttggttgacaag 1320
Db 1261 ttggttagattttggtttaacgacagagtttccattcgcacggttggttggttgacaag 1320
Qy 1321 ttggttagattttggtttaacgacagagtttccattcgcacggttggttggttgacaag 1380
Db 1321 ttggttagattttggtttaacgacagagtttccattcgcacggttggttggttgacaag 1380
Qy 1381 aactggggaagaagtgttctcctta 1403
Db 1381 aactggggaagaagtgttctcctta 1403

```

RESULT 5

US-08-419-448-33
 Sequence 33, Application US/08419448
 Patent No. 5863533
 GENERAL INFORMATION:
 APPLICANT: Robert F.M. Van Gorcom
 APPLICANT: Willem Van Hartingsveldt
 APPLICANT: Petrus A. Van Paridon
 APPLICANT: Annemarie E. Veenstra
 APPLICANT: Rudolf G.M. Luttin
 APPLICANT: Gerardus Seiten
 TITLE OF INVENTION: Cloning and Expression of Microbial
 TITLE OF INVENTION: Phytase
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006-1888
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 SOFTWARE:
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/419,448
 FILING DATE: 10-APR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Murashige, Kate H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 24615-20026.10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-887-1500
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:

100

;	LENGTH: 1404 base pairs	
;	TYPE: nucleic acid.	
;	STRANDEDNESS: double	
;	TOPOLOGY: linear	
;	MOLECULE TYPE: cDNA to mRNA	
;	HYPOTHETICAL: NO	
;	ANTI-SENSE: NO	
;	ORIGINAL SOURCE:	
;	ORGANISM: <i>Aspergillus ficuum</i> (<i>Aspergillus niger</i>)	
;	STRAIN: NRRL 3135	
;	US-08-419-448-33	
	Query Match 43.7%; Score 614.2; DB 2; Length 1404;	
	Best Local Similarity 64.9%; Pred. No. 4.6e-173;	
	Matches 910; Conservative 0; Mismatches 493; Indels 0; Gaps	
QY	1 atggcgctgtctgctgactgtccattgcccactgttctgcgttccacatccggtacc	60
Db	1 ATGGCGCTGCTCGTGTCTACTTCCCTTTGTATCTCCGTCTGGAGTCACTCCGGACTG	60
QY	61 gctttgggtctctggttaactctactctgtgacactgttgacggtgggtaccaatgt	120
Db	61 GCAGTCCCGCCCTCGAGAAATCAATCCAGTTCGGATACGGTCGATCAGGGGTATCAATGC	120
QY	121 ttcccagaataattctcacttgggtgggtaactctcattctctctcttgggtgacgaa	180
Db	121 TTCTCCGAGACTTCGGCATCTTTGGGTCAATATGCGACCGCTTCTTCTCTCTGGCAACGAA	180
QY	181 tctctatttctccagacttccaaaggtgtgagttactcttccttcgaatttgcct	240
Db	181 TCGGTCTATCTCCCTGAGTGGCCCGCGGATCCAGATGACTTTTCGCTCAGGTCCTCTCC	240
QY	241 agacacggtgctagataccacaacttctctgctgaaggctactctgctttgattgaa	300
Db	241 CGTCATGGAGCGGGTATCGACCGACTCCAAGGCCAAGAAATACTCCGCTCTCATTTGAG	300
QY	301 gctattcaagaacgctactgctttcaagggttaagtaacgctttcttgaagacttaaac	360
Db	301 GAGATCCAGCAGAACGCGACACCTTTTGACGAAATATGCTTTCCTGAAGACATACAAC	360
QY	361 tacactttgggtgtgacgacttgaactcattcggtgaacacaaatggttaactctgt	420
Db	361 TACAGCTTGGGTGCAGATGACTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGC	420
QY	421 ataatgtctacagaagatacaagggtttggctagaagaattgttccattcattagact	480
Db	421 ATCAAGTTCTACCAAGCGGTACGAATCGCTCACAAAGGAACATCGTTCCATTTATCCGATCC	480
QY	481 tctggttctgacagagtattgcttctgctgaaaagttcaatgaaaggtttccaatctgt	540
Db	481 TCTGGCTCCAGCCGCGTATPCGCTCCGCAAGAAATTCATCGAGGGCTTCAGAGCACCC	540
QY	541 aagttggctgaccaggtgtcaaccacacaaagcttctccagttataagcttattatt	600
Db	541 AAGCTGAAGATCTCTCGTGCCCGCCGCCGCAATCGTCCGCCAAGATCGAGCTGGTGCATT	600
QY	601 ccagaaggctcgtggttacacaacactttggaccacggtttgtgtactgcttttcgaagaa	660
Db	601 TCCGAGGCCAGCTCATCCAACAACACTCTCGACCCAGGCACCTGCACCTGTCTTCGAAGAC	660
QY	661 tctgaattgggtgacgactgaagctaaactcaactcgtgttttcgctccacaattaga	720
Db	661 AGCGAAATGGCCGATACCGTFCGAAGCAATTTACCGCCACGTTTCGTCCCTCCATTCGT	720
QY	721 gctagattggaagctcaacttgcagaggtttaacttgactgacgaagcgttgtttaactg	780
Db	721 CAACGCTCTGGAGAACGACCTGTCCGCTGTGACTCTCACAGACACAGAGTGACCTACCTC	780
QY	781 atggacatgtgtccattcgacactgttgttagaactcttgacgctactcaattgtccaa	840
Db	781 ATGACATGTGCTCTTCGACACCATCTCCACGACACCGTCGACACCAAGCTGTCCCCG	840

Qy	841	tctctgacttgtcaactcaagacgaatggattcaatacgaactacttgcaattcttgggt	900
Db	841	TTCTGTGACCTTTTTTCCACCATCAGCAATGGATTAACTACGACTACCTCCAAGTCCTTGAAA	900
Qy	901	aagtactcggttacgctgcgtgtaaacccattgggtccaagtciaagtgctgttclcggt	960
Db	901	AAGTATTACGGGCATGGTGCAGGTAAACCCGCTCGGCCGACCCAGGCGTTCGGCTAGGCT	960
Qy	961	aacgaattgattgttagattgacctactctocagttcaagaaccacacttctactaaaccac	1020
Db	961	AACGAGCTCATGCGCCGCTGCAGCACCTCGCCTCTCCACGATGACACCAGTTCACAACCAC	1020
Qy	1021	aatttggacttaaccccagctactttcccatgaaocgctacttttgttaoigtgactctetct	1080
Db	1021	ACTTTTGGACTTCAGCGCCGGCTACCTTTTCGGTCAACTCTACTCTCTACGCGGAATTTTCG	1080
Qy	1081	cacgacaacactatggtttctatttctctgcctttgggtttgtgtacaacgtactaaagcca	1140
Db	1081	CATGACACGGGATCATCTCCATCTCTTTGCTTTAGGCTGTGTAACAGGCACCTAAGCCG	1140
Qy	1141	tgtctactactcttcttgaaactattgaagaaactcacggttactctgtctcttggact	1200
Db	1141	CTATCTACCAACACCGTGGAGAATAATACCCACAGACATGATTCGTCTGCTTGGACG	1200
Qy	1201	gttcattctgcgtctagagcttacgttgaalgtatgcaatgtgaagctgaaaaggaaacca	1260
Db	1201	GTTCCGCTTTGCTTCGCGTTTTCAGCTCAGATGATGCAGTGTCAAGCGGAGCAGGAGCCG	1260
Qy	1261	tgggttagagtttgggttaaogacagagattgttcattgcacggttgtgtgtgtgacaag	1320
Db	1261	CTGGCCCGTGCTTGGTTAAATGATCGCGTTGTCCCGCTGCAATGGGTGTCCGTTGAFGCT	1320
Qy	1321	tgggttagatgttaagagagacgacttcgttgaaggtttgtcttttcgatagatctgggtgt	1380
Db	1321	TTGGGGAGATGTACCCGGATAGCTTTGTGAGGGGTTGACCTTGTGATAGTCTGGGGGT	1380
Qy	1381	aactgggaagaatgtttcgccta	1403
Db	1381	GATTGGCGGAGTGTTTTGCCTTA	1403

```

RESULT 6
US-07-923-724-7
; Sequence 7, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloheimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Eagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992

```

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/496,155
 FILING DATE: 19-MAR-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/044,077
 FILING DATE: 29-APR-1987
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 8610600
 FILING DATE: 30-APR-1986
 ATTORNEY/AGENT INFORMATION:
 NAME: Cimbalà, Michele A.
 REGISTRATION NUMBER: 33,851
 REFERENCE/DOCKET NUMBER: 1050.0240004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2363 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: both
 FEATURE:
 NAME/KEY: CDS
 LOCATION: join(404..447, 550..1906)
 US-07-923-724-7

Query Match 43.5%; Score 610.2; DB 1; Length 2363;
 Best Local Similarity 64.9%; Pred. No. 9e-172;
 Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Qy 13 gtcgtgactgtcattgacacactgttgcgttccacatcggtacgcgttggtgct 72
 Db 518 GTGGGACTACTGATCGCTGCAATCTGTGCAGAGTCACTCCGGACTCGGCTCCCGCC 577
 Qy 73 cgtgtaactctactctgtgacactgttgacggtggttaccactgtttccagaaatt 132
 Db 578 TCGAGAAATCAATCACCTTGCATAGGTGCGATCAAGGTATCAATGCTTCCGAGACT 637
 Qy 133 tctcaactgtgggtacatactctccatctctctctctctctctctctctctctctct 192
 Db 638 TCGGACTCTTGGGGTCAATCCCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 697
 Qy 193 ccagacgttccaaagggtgtgtaggttactctgttgcgttactctgttgcgttactct 252
 Db 698 CCTGATGTGCCCGCGGTGTGCAGAGTCACTTTCGCTCAGGTCTCTCTCCGCTCATGGAGCG 757
 Qy 253 agatacccaacttctctgctctaaaggcgtactctgttgcgttactctgttgcgttactct 312
 Db 758 CGGTATCCGACCGAGTCCCAAGGGCAAGAAATVACTCCGCTCTCATTTAGGAGATCCAGCAG 817
 Qy 313 aacgactgtcttcaagggttaagtaagcttcttgaagacttcaactacacttgggt 372
 Db 818 AACGTGACCACTTTGATGGAATAATATGCCCTCTCTGTAAGACATCAACTACAGCTTGGGT 877
 Qy 373 gctgacgacttgactccattcgttgaaacaaatggttgaactctgttgaattactac 432
 Db 878 GCAGATGACCTGACTCCCTTCGGAGAGCAGGAGTACTCAACTCCGGCATCAAGTTCTAC 937
 Qy 433 agaagatacaagggttggctagaaagtgttccattcattagacttctggttctgac 492
 Db 938 CAGCGATACGAATCGCTCACAAGGAACATCATTCGGTTTCACTCCGATCTCTGGCTCCAGC 997
 Qy 493 agacttattctctgtgaaagtcttgaaggtttccaatctgtaagtgtggtgac 552
 Db 998 CGCGTATCGCCTCCGGGGAATTCATTGAGGGCTTCTAGAGACCAACCAAGGAT 1057
 Qy 553 ccaggtgttaacccacacacagcttctccagttattacgttattattccagaaggtgct 612
 Db 1058 CCTCGTCCCGCGGCGCAATCTCGCCCAAGATCGACGTGGTCATTTCCGAGGCCAGC 1117

Qy 613 ggttaacaaacacacttggaccacggttctgttactgcttctgaagaatctgaattgggt 672
 Db 1118 TCATCCACACACACTCTCGACCCAGGCACCTGCATCTGTTTGAAGACAGCAATTTGGGC 1177
 Qy 673 gacgagcttgaagctaaacttcaactgcttttgcgtccaccacattagactgagattggaa 732
 Db 1178 GATACCGTCGAAGCAATTTCCACCGCCAGGTTCGCCCTCCATTCGTCAACGCTGGAG 1237
 Qy 733 gctcaactgccagggtttaacttgaactgacgaagcgttggtaacttggatgacatgtgt 792
 Db 1238 AACGACCTGCTGGGCTGACTCTCAAGACACAGAAAGTACCTACCTCATGGACATGTGC 1297
 Qy 793 caattgacactgttgcctagaaacttcaagcgtactactaaatttctccattctgtgacttg 852
 Db 1298 TCCCTTCGACACCATCTCCACAGCAGCCGTCGACACCAAGCTGTCCCTCTCTGTGACCTG 1357
 Qy 853 ttactcagcagaatggattcaatcacgactacttgcgaatctttggtaagtactacgt 912
 Db 1358 TTCACCCATGACGAATGGATCCACTACGACTACCTCCAGTCCCTGANAATACTACGGC 1417
 Qy 913 tacggtgctgttaacccattgggtccagctcgaagtgttggtttgcgttaacgaattgatt 972
 Db 1418 CATGCCGAGGTAAACCGCTCGGCCGAGCCAGGCGTGGCTACGCTAACGAGCTCATC 1477
 Qy 973 gctagatgactcactctccagttcaaacacacacttctactaaacacacttggactct 1032
 Db 1478 GCCGCTCACCCACTCGGCTGTCCACGATGACACAGCTCCCAACACACACTTGGACTCG 1537
 Qy 1033 aaccagactcttccacttgaacgctacttctgtacgctacttctctcaacacacact 1092
 Db 1538 AACCCAGCTACTCTCCGCTCAACTCTACTCTCTAGCGGACTTTTCCACGATACGGC 1597
 Qy 1093 atggttctatttcttgcgttgggtttgttaacacggttactaaagccattgtctactact 1152
 Db 1598 ATCAFTCTATCTCTCTTGTGGTCTGTACAACGGCAGCTAACCGCGTGTCTACCAAG 1657
 Qy 1153 tctgtgactctatgaagaactgacggttactctgttctgttgcgttctccactgct 1212
 Db 1658 ACCGTGGAGATATACCCAGACAGATGGGTCTCTGCTGTGTGGACGGTTCCTGTTGCT 1717
 Qy 1213 gctagagcttactgttgaatgatgcaatgtgaagctgaaaggaacccattggttagagt 1272
 Db 1718 TCGCTCTGTACGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1777
 Qy 1273 ttggttaacacagaggttgccttccattgacggttgcgttgcgttgcgttgcgttgcgtt 1332
 Db 1778 TTGGTTAATGATCGGCTTGTCCCGCTGCATGGGTGTCCAAATGTGATGCTTTGGGGAGATG 1837
 Qy 1333 aagagagacgacttgcgttgaaggttgcgttcttctgcgtactgctggttaactgggaagaa 1392
 Db 1838 ACCCGGATAGCTTTGTGAGGGGTTGAGCTTTGCTAGATCTGGGGGTCATTTGGCGGAG 1897
 Qy 1393 tgttctgctta 1403
 Db 1898 TCTTCTGCTTA 1908

RESULT 7
 US-08-609-426A-7
 ; Sequence 7, Application US/08609426A
 ; Patent No. 5830733
 ; GENERAL INFORMATION:
 ; APPLICANT: Nevalainen, Helena K.M.
 ; APPLICANT: Palheimo, Marja T.
 ; APPLICANT: Miettinen-Oinonen, Arja S.K.
 ; APPLICANT: Torikeli, Tuula K.
 ; APPLICANT: Cantrell, Michael
 ; APPLICANT: Piddington, Christopher S.
 ; APPLICANT: Rambosk, John A.
 ; APPLICANT: Turunen, Marja K.
 ; APPLICANT: Fagerstr m, Richard B.
 ; APPLICANT: Houston, Christine S.
 ; TITLE OF INVENTION: Production of Phytase Degrading Enzymes

Db 864 COTCGTGGCCACCGCGCAATCTGCGCCAAAGATCGAGTGGTCTATTCCGAGGCCAGC 923
Qy 613 ggttaacaacacactttggaccacggtttgtgtactgtcttccgaagaatctgaattggt 672
Db 924 TCATCCACACACACTCTCGACCCAGCCAGCTGCTGCTCGAAGACAGCGAATTGGCC 983
Qy 673 gacgactggaagctaaacttactgtcttttcgtccccaacttagagctagattgaa 732
Db 984 GATACCGTGAAGCAATTCACCCCGAGCTGCTGCTCCCTCCATTCGTAACGCTGGAG 1043
Qy 733 gctcaactgcaggtgttaacttactgacgaagacggttgaacttgaatgacatgtgt 792
Db 1044 AACGACCTGTCCGCTGACTCTACACACACAGAGTACCTCATGTGACATGTC 1103
Qy 793 ccattcgacactgttgtagaactctacgctactcaactgtctccattctgtgacttg 852
Db 1104 TCCTTCGACACCATCTCCACAGACCGTCGACACCAAGCTGTCCCTCTGTGACCTG 1163
Qy 853 ttactcagcgaatgattcaatacactacttgcacttcttgggttaagtactacggt 912
Db 1164 TTCACCCATGAGGATGATCAACTACGACTACCTCCAGTCTTGAAGAAGTATTACGGC 1223
Qy 913 tacggtgtgttaacacacttgggtccagctcaagtggtgttgcgttaacgaattgatt 972
Db 1224 CATGTGTCAGGTAACCGCTCGCCCGACCCAGGGCGTGGCTAGCTAACGAGCTCATC 1283
Qy 973 gctagatgactcactctccagttcaagacacacacttctactaacacacacttgactct 1032
Db 1284 GCCGTGTGACCATCTCGCTGTCACGATGACACCAAGTTCACCAACACACTTTGGAGCTG 1343
Qy 1033 aaccagctacttcccaattgaacgtacttctgtacgtgacttctctcacacacact 1092
Db 1344 AGCCCGGTACTCTTCCGCTCAACTACTCTCTACGGGACTTTTCGATGACACAGGC 1403
Qy 1093 atggtttctatttcttcgtcttgggttgaacacggttaacgaactgcttactact 1152
Db 1404 ATCATCTCCATCTCTTGTCTGTAGTGTGTACACGCGCACTAAGCGGTATCTACCA 1463
Qy 1153 tctgtgaactattgaagaactcaaggttactctgcttctgtgactgttccattcgt 1212
Db 1464 ACCGTGGAGAAATCACCCAGACAGATGATCTCTGCTGTGGACGGTTCGGTTGCT 1523
Qy 1213 gctagagcttactgtgaatgatgcaatgtgaagctgaagaaagaccattggttagatt 1272
Db 1524 TCGCGTTGTAGCTCGAGATGATGAGTGTACAGCGGACGACGAGCGCTGTCGCTGTC 1583
Qy 1273 ttggttaacgacagattgttccattgcaacggttgggtgttgacaaagtgggttagatt 1332
Db 1584 TTGGTTAATGATCGCGTTGTCTCCCGTGCATGGGTGTCGCGTTGATGCTTTGGGAGATGT 1643
Qy 1333 aagagagacgacttctgtgaaggttcttcttcgctagatctcgttggttaactgggaagaa 1392
Db 1644 ACCCGGATAGCTTGTGTGAGGGGGTTGACCTTTGTAGATCTGGGGGTGATTGGCGGAG 1703
Qy 1393 tgttcgctta 1403
Db 1704 TGTGTTGCTTA 1714

RESULT 10

US-08-419-448-31
; Sequence 31, Application US/08419448
; Patent No. 5863533
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gorkom
; APPLICANT: Willem Van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Seiten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; IMMEDIATE SOURCE:
; LIBRARY: lambda AF
; CLONE: pAF2-3, pAF2-6, pAF2-7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 210..253
; FEATURE:
; NAME/KEY: intron
; LOCATION: 254..355
; FEATURE:
; NAME/KEY: exon
; LOCATION: 356..1715
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(210..253, 356..1715)
; OTHER INFORMATION: /codon_start=210
; -OTHER INFORMATION: /product="Phytase"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 210..380
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 381..1712
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function="inositol phosphate
; OTHER INFORMATION: phosphatase"
; OTHER INFORMATION: /product="Phytase"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; US-08-419-448-31

Query Match 43.5%; Score 610.2; DB 2; Length 6756;
Best Local Similarity 64.9%; Pred. No. 1.5e-171;
Matches 903; Conservative 0; Mismatches 488; Indels 0; Gaps 0;

Oy 13 gtcgtgtactgtccattgcccacttgcgttccacatccggtaccgcttgggtcct 72
Db 324 GTGGGACTACTGATCGCTCACTATCTGTCGAGAGTCACTCCGAGTCCGAGTCCCGGCC 383

Qy	73	cgtagtaactctcaactcttggacacactgttgacggtggttaccaaatgtttccagaaatt	132
Db	384	TCGAGAATCAATCCAGTTCGGATACGGTCGATACAGGGGTATCAATGCTTCTCCGAGACT	443
Qy	133	tctcaacttgggggtacatactctccatctctctcttggctgagcaaatctgctatttt	192
Db	444	TCGCATCTTTGGGGTCAATACGACCAGTTCCTCTCTGGCAACGAATCGGTCACTCC	503
Qy	193	ccagcgttccaaagggttgtaggttaactcttggttcaaagtttctctagacaagtgct	252
Db	504	CCTGAGTGC CGCGGATCGAGGTCACTTTGCTCAGGTCTCTCCCGTCATGGACG	563
Qy	253	agatacccaactttcttgogtctaaggogtactctbgtttgattgaagctattccaaag	312
Db	564	CGGTATCCGACCGACTCCAAAGGCAAGAATACTCCGCTCTCATTTAGGAGATCCAGCAG	623
Qy	313	aacgctactgcttcaaagggtaagtacgctctctcttgaagacttcaactacactttgggt	372
Db	624	AACGGACCAACOTTGACGGAAATATGCTTCTCTGAAGACATACAACTACAGCTTGGGT	683
Qy	373	gctgaogacttgactccttcggtgaacacaaatgggttaactctggtattaaagtctctac	432
Db	684	GCAGATGACCTGACTCCCTTCGGAGAACAGGAGCTAGTCAACTCCGGCATCAAGTTCTAC	743
Qy	433	agaagatacaaggctttgctagaagaattgttccattcattagagcttctgggtctgac	492
Db	744	CAGCGGTACGAATCCGCTCACAGGAACATCGTGTCCATTCATCCGATCTCTGGCTCCAGC	803
Qy	493	agagttattgctctgctgaaaagttcattgaaggtttccaatctgctaaagtggctgac	552
Db	804	CGCGTGATCGCCTCGGCAAGAAATTCATCAGGGGCTCCAGAGCACCAAGCTGAAGGAT	863
Qy	553	ccagggtgaacccacacaaagcttccagttattaaagttattattccagaagagtgct	612
Db	864	CCTGTGCCACGACCGCGCCAAATCGTCGCCAAGATCGACGTGGTCAATTCGGAGGCCAGC	923
Qy	613	ggttacaacaacactttgaccacagtttgttactgcttctcagaagaatctgaattgggt	672
Db	924	TCATCCAAACAACTCTCGACCCAGGACCTGCATCTGCTTCGAAAGACGGCAATGGCC	983
Qy	673	gacgagttgaagctaacttcaactgctgttttcgctccaccaataggagctagattggaa	732
Db	984	GATACGCTGAAGCGAATTCACCGCCAGTTGCTGCTCCCTCCATTCGTCAAGCTGGAG	1043
Qy	733	gctcaactggcaggtgttaacttgactgacgaagacgttgttaacttgatggacatgtgt	792
Db	1044	AACGACCTGTCCGGTGTGACTCTCAGACACAGAAAGTGACTACCTCATGGACATGTC	1103
Qy	793	ccattgcaactgtgctagaactctgacgtactcaatgtctcaattctgtaattg	852
Db	1104	TCCTTCGACACCATCTCCACGACCGTCGACACCAAGTGTCCCTCTCTGTGACCTG	1163
Qy	853	ttcactcacgacgaatggattcaatacgaactacttgcactcttgggttaagtactacggt	912
Db	1164	TTCAACCATGACGAATGGATCAACTACGACTACCTCCAGTCTCTGAAAAAGTATTACGGC	1223
Qy	913	taagggtgctggtaaccocatgggtccagctcaaggtgttggtttgttaacgaattgatt	972
Db	1224	CATGGTGCAGGTAAACCGCTCGCGCCGACCCAGSGCGTCGGCTACGTTAAGAGCTCATC	1283
Qy	973	gctagattgactcaactctccagttcaagacacacttctactaaccacactttgactct	1032
Db	1284	GCCGCTGNACCATCTGGCTGTCCACGATGACACCAAGTTCACACCACTTTGGACTCG	1343
Qy	1033	aaccagctactttcccatgaaacgtactttgttaagctgacttctctcacgaacact	1092
Db	1344	AGCCCGGCTACCTTTCCGCTCAACTCTACTCTCTACGGGACTTTTCGATGACACGGC	1403
Qy	1093	atggttctattttcttgcgtttgggtttgtacaacggtactaagccatttgtactact	1152
Db	1404	ATCATCTCAATCTCTTTGCTTTAGGCTGTGACACGCGCACTAAGCGGCTATCTACCACG	1463

Qy	1153	tctgtgaattctattaaagaaactgcacggttactctgttcttcttggaactgttccattcgcgt	12112
Db	1464	ACCCTGGAGAAATATCACCCAGACAGATGCTCTGCTTTGGACGCTTTCCGTTTGCT	15233
Qy	1213	gctagagcttaccttgaattgacgaatgtgaagctgaaaggaacattgttttagagtt	12772
Db	1524	TCGCGTTTGTACGTCGAGATGATCGACGCTCAGGCGGACGACGCCCTCGTCCGTTGC	15833
Qy	1273	ttggtttaacgacagagtgtttccattgcacggttctggtgtttgacaaagttgggttagatgt	13322
Db	1584	TTGCTTTAATGATCCGCTTGCTCCCGCTGCATGGGTCTCCGTTGTATGCTTTGGGAGATGT	16433
Qy	1333	aagagagacacacttcgttgaaggtttgtcttcttcgtagatctggttgtaactgggaagaa	13922
Db	1644	ACCCGGGATAGCTTTGTGAGGGGGTTGACGTTTGTAGATCTGGGGGTGATTGGCGGGAG	17033
Qy	1393	tgtttcgctta	14033
Db	1704	TGTTTTCGCTTA	17143

RESULT 11

```

US-09-155-855-5
: Sequence 5, Application US/09155855
: Patent No. 6139902
: GENERAL INFORMATION:
: APPLICANT: KONDO, Hiemasa
: APPLICANT: ANAZAWA, Hideharu
: APPLICANT: KANEKO, Syunichi
: APPLICANT: NAGASHIMA, Tadashi
: APPLICANT: TANGE, Tatsuya
: TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
: FILE REFERENCE: 81356/124
: CURRENT APPLICATION NUMBER: US/09/155.855
: CURRENT FILING DATE: 1998-10-05
: EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
: EARLIER FILING DATE: 1997-04-04
: EARLIER APPLICATION NUMBER: JP 084314
: EARLIER FILING DATE: 1996-04-05
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 5
: LENGTH: 1515
: TYPE: DNA
: ORGANISM: Aspergillus niger
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: (1)..(45)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(45)
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: (157)..(183)
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: (184)..(1512)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (157)..(1512)
:
: US-09-155-855-5

```

	Query Match	42.1%	score 591.6;	DB 3;	Length 1515;
	Best Local Similarity	64.6%;	Prod. No. 2.5e-166;		
	Matches	882;	Conservative	0;	Mismatches 484;
					Indels 0;
					Gaps 0;
QY	38	tgttcggttcacacatccggtacccgcttgggtctctcgttggtgaactctcacctctctgtgacac	97		
Db	149	tgcacagatcacctccggtacgtggcgtcccgctccgagaaatccaatccacttcgcata	208		
QY	98	ctgtttgacggtgggttaccacatgtttccacagaaattctcacttgggtggggtcacactctc	157		

Db 209 cggctgacaggggtatcaatgtctctcggagacttcgcatcttggggccaaacacgcgc 268
Qy 158 cattctctcttgctgacgaatctgtatttctccagaagttccaaagggttttagag 217
Db 269 cgttctctcttgccaaacaaatcggcatctccctctgatttctctcggatgccatg 328
Qy 218 ttaacttctgacaaatgtttctagacaggtgtctagatcccaacttctctcgtctta 277
Db 329 tcaacttccgacaggtctctccgcacatggagacggtatccgacccgactccaaaggcca 388
Qy 278 aggcgtactctgtttgattgaagctatccaaagaacgctactcttccaaagggttaagt 337
Db 389 agaaatactccgctctcatcctgagggatccagcaagacgcgacaaccttcgaggggaaat 448
Qy 338 acgctttctgaagaacttaacatacactatcttgggtgtctgagacgttgaactccatcggtg 397
Db 449 atgctctctgaagacatacaactacagactcggcgcgagatgactgactccctcggag 508
Qy 398 aacaacaaatggttaactctgtatttaagtctacagaagatacaaggcttttggctagaa 457
Db 509 agcagagagctgtcaactccggcgctcaagttctaccagcgatacgaatcgctcaacaagaa 568
Qy 458 agattgtccattcattagagctctgtgtgtctgacagagttattgtctctgtctgaaaagt 517
Db 569 acattgtccgcttcatcgtactcaggctccagcgcggtgattgctctgtgcaataaat 628
Qy 518 tcaattgaaggtttccaaatctgtaagtgtgtgacccaggtgtactaaccacacaaagctt 577
Db 629 tcatcgagggctccagacactaagctgaaggtactcctgtgccagccggcccaatgt 688
Qy 578 ctccaggtatttaacgttattatccagaaggtgtgtgttaccacacacacttggaccacg 637
Db 689 cggccaaagatcggtgtcatttcagaggcgacacacacacacactctcgatccgg 748
Qy 638 gtttgtactgtcttccgaagaatctgaattgggtgacgacgttgaagtaacttcaactg 697
Db 749 gcacctgcacctgtttgaagatagcgaattgtgacgtgacatcgaaagcgaatttcaacg 808
Qy 698 ctgtttcgtcccaacttagatagatgtgaagctcaattgcccaggtgttaacttga 757
Db 809 ccacgttctcctccatctcactcagctcgtggaagacgttctggtgctgtctca 868
Qy 758 ctgacgaagacgtgttaactgtatggacatgtgtccattcgacacgtgtgtgagaactt 817
Db 869 cggacacagaagtgacctaccctcgtgacatgtgtctctcgacacacactccaccagca 928
Qy 818 ctgacgctactcaattgtctcattctgtgacttgttcaactcagcagcaatggattcaat 877
Db 929 ccgtcgacacacagctgtccctctctgtgacctgttccacctgaagaatggatcaact 988
Qy 878 acgactacttgaacttttgggtgaagtaactacggttactcgtgtgtgtaaccacttgggtc 937
Db 989 acgactactcagctccttgaacaaataactacggtcgtggaagacgttctggtgctctca 1048
Qy 938 cagctcaaggtgtgtgttctgtaacgaatgtatgtctgactgactcactctccagttc 997
Db 1049 cagccagggtcgtgctacgtacgtacgtcgtcccgctcaccactcgtctgtctc 1108
Qy 998 agagacacacttctactaaccacacttttggacttcaaccagactacttcccaatgaacg 1057
Db 1109 acgatgacacagctcccaacacacttggacttcccaaccggctacttctccgctcaact 1168
Qy 1058 ctacttgtacgctgactctctcagcaacacactatggtttctctatttctcgtcttgg 1117
Db 1169 ccactctctatcggaactttcgtatgaacggcatctctctatctctctcttctgttgg 1228
Qy 1118 gtttgtacacggttactaagcattgtctactacttctgttgaattattgaagaactg 1177
Db 1229 gctgtacacagcgacacagcgctgtcttccacgacccgagagaataccaccagacgc 1288
Qy 1178 acggttactctctcttctgactgttccattcgtctgtagacgttactcgttgaatgatgc 1237
Db 1289 atgggttctcactcgtcggcgggttctccttctcgtcgtcgcgtacgtcgcagatgatgc 1348

Qy 1238 aatgtgaagctgaaaggaagcaaccattggttagagtttgggttaacagacagagttgttccat 1297
Db 1349 aatgcccagctccagagcagagcccttggctccggtgtcttgggttaatgactgtgttccgc 1408
Qy 1298 tgcacggttgggtgttgcacaagttgggttagatgttaagagagacgactcgtttgaaggtt 1357
Db 1409 tgcattggtgtccggttgcattgttgggaagatgtacgcgggagtagcttcgtgaagggtt 1468
Qy 1358 tgcatttgcgtagatcgtgtggttaactgggaagaatgttttcgctta 1403
Db 1469 tgcatttgcgtagatcgtgtggtggtggtggtggtggtggtggtggtggttcttctta 1514

RESULT 12
US-09-155-855-4
; Sequence 4, Application US/09155855
; Patent No. 6139902
; GENERAL INFORMATION:
; APPLICANT: KONDO, Hidemasa
; APPLICANT: ANAZAWA, Hideharu
; APPLICANT: KANAKO, Syunichi
; APPLICANT: NAGASHIMA, Tadashi
; APPLICANT: TANGE, Tatsuya
; TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
; FILE REFERENCE: 81356/124
; CURRENT APPLICATION NUMBER: US/09/155,855
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
; EARLIER FILING DATE: 1997-04-04
; EARLIER APPLICATION NUMBER: JP 084314
; EARLIER FILING DATE: 1996-04-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1332)
US-09-155-855-4

Query Match 42.1%; Score 590.6; DB 3; Length 1332;
Best Local Similarity 65.5%; Pred. No. 4, 6e-166;
Matches 863; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

Qy 87 ctcttgtgacactgttgacggtgttaccatgtttccagaaaatttctcaacttggggg 146
Db 15 caactgtcgatacggctcgatcaggggtatcaatgtcttcggagactcgcatttggggg 74
Qy 147 tacatactctccattctctcttctgtgctgacgaatctgtatttctccagacgttccaaa 206
Db 75 ccaatacgcgcgttcttctctgtgcaacaaatcgccatctccctcgtatgtctgc 134
Qy 207 ggggtgtagagttacttctgttcaagttttgttagacacggtgtgtagatacccaacttc 266
Db 135 cggatgccatgtcatttctgcgcaggttctctcccgatggagcgggtatccgacga 194
Qy 267 tctcgcgttaaggcgtactctgtcttgcattgaagctatttcaaaaagaacgctactgctt 326
Db 195 ctcaaggcgaagaataactcgcctctcctcgcagagatccacagacagcgcgaacactt 254
Qy 327 caagggtgaagtaacgcttttcttggaagacttacaactacaacttgggtgtcgtacgacttgac 386
Db 255 cgaggggaaatagcttctcctgaagacatacaactacagcctggtggtggtggtgac 314
Qy 387 tccattcgtggaacacaaatggttaactctgttatttaagtttctacagaagatacaagc 446
Db 315 tcccttcgagagcagaggtgtgtcactccgcgtcaagttctcaccagcgaacgaatc 374
Qy 447 ttgggttagaagaatgttcttcattcattagagcttctgttctgtgacagagttattgtcttc 506

```

Db 375 gctcacaagaacatgtccgttccatcgatctccaggtccagccggtgatgctc 434
QY 507 tgcgtgaaggttcattgaaggtttccatctgctgaagttggtgacccaggtgctaaacc 566
Db 435 tggcaataaattcatcgaggttcagagcagatgctgaaggtatctctggtccagcc 494
QY 567 acaccaagcttccagttatataaagttatttccagaaggtgctggtttacacaaacac 526
Db 495 cggccaatcgtcccaagatcgacgtggtctatcttcagagggccagccacatccacaaacac 554
QY 627 ttggaccacgggttggtagctgcttgaagaatctgaattgggtgacgacggttgaagc 686
Db 555 tctgcacgggacacgtcaccgttttccagatagcgaattggcgtgacatcgagcagc 614
QY 687 taactcactgctgttttcgtccacaaattagagctagattggaagctcacttgcacgg 746
Db 615 caatttcacgcacgttctgctccctccatctgccaagctctggagaacgacttctcgg 674
QY 747 tgttaacttgactgacgagacgtgttaacttgatgacatgtgtccattcgacactgt 806
Db 675 cgtgtctcagcagacagagtgacacctacctcatgagacatggtctccttcgacacac 734
QY 807 tgtagaactctgacgtactcaattgtctccattctgtgacttctgaactcactcagcagc 866
Db 735 ctccaccagcagctcgacacacagctgcccctcttctgacactgttcacccatgaaga 794
QY 867 atgattcaatacgaactacttgaactcttctggttgagtagctacggttgaagctggttaa 926
Db 795 atggatacaactgactacactcctcagctccctgacacaaatctactacggtccatggc 854
QY 927 cccattgggtccagctcaaggttgggttctggttaacgaattgattgctagattgaactca 986
Db 855 ccgctcggccgacccagggcgtcggtctacgtcactacagctcactcgcctctcaccaca 914
QY 987 ctctccagttcagaacacacttctactaaccacacacttctgactctaacccagctacttt 1046
Db 915 ctgcctgtccacgatgacacacagctcccaacacacacattggactcccaaccggctacttt 974
QY 1047 cccattgaagctacttctgacgtgaacttctctcactcactcactcactcactcactt 1106
Db 975 ccgctcactcactcactcactcactcactcactcactcactcactcactcactcactcact 1034
QY 1107 cttcgtttgggttgaacacggtactaagcacttctcactcactcactcactcactcactcact 1166
Db 1035 ctttgggttgggttgaacacggtactaagcactcactcactcactcactcactcactcact 1094
QY 1167 tgaagaactgacggttactctgttcttggactgttccattcgtcgtgactgactgact 1226
Db 1095 caccacagaccggtgtctcactcgtcgtgacgttcttctcgtcgtcgtcgtcgtcgtcgt 1154
QY 1227 tgaatgatgaatgtgaagctgaaaggaagacccattggttagaggtttggttaacgacag 1286
Db 1155 cgagatgatgaatgccagtcgacgagcagcgttcttgcgtgttcttgcgtgttgcgtgttgc 1214
QY 1287 agttgtccatgacggttctggtgttgacaaagttggttagatgtagagagacgactt 1346
Db 1215 tgttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1274
QY 1347 cgttgaaggttcttctcgttagatctggttgaactgggaagaaagtcttcgctta 1403
Db 1275 cgtgaaggggttgagcttgcagatctggtcgtggtggtggtggtggtggtggtggtggt 1331

```

RESULT 13

US-08-819-825-1

; Sequence 1, Application US/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 4758.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2200 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-819-825-1

Query Match 24.6%; Score 346; DB 2; Length 2200;

Best Local Similarity 57.2%; Pred. No. 3 le-93;

Matches 753; Conservative 0; Mismatches 515; Indels 48; Gaps 5;

```

QY 127 gaaattctcaacttgggtggtacatactctccattcttcttggctgacgaattgct 186
Db 607 GACATTCCGCCCACTGGGCCAGTACTGCCCTTCTTCGCTGGCCGAGGCTCTGAA 666
QY 187 atttctccagacttccaaagggttagagttacttctggttcaagtttctgctagacac 246
Db 667 ATCTCGCTCGGTGGCCCAAGGGCTGTCTGTCGAGTGTTCGACAGTGTCTGCCGCCAC 726
QY 247 ggtgctagataccacacttctctgctgcttaagcgtactctgcttggattgaagctatt 306
Db 727 GGAGCTCGGTATCTCTACTGCTCACAGAGTGAAGTCTACGCCGAGTTGCTTCNAAGGATC 786
QY 307 caaagaacgctactgctttcaagggttaagtagcgttcttctgaagacttacaactacact 366
Db 787 CAGGACACTCGCAGCGAGTTCAAGGGCGATTTTGCCTTCTCCGAGACTATGCCTATCAT 846
QY 367 ttgggtgctgacactgactcctcattcgttggaacaaataagtttaactctggtttaag 426
Db 847 CTCGGTCCCGATTAATTGTCGCGCTTTTGGCGAGGACACAGATGATGGAATCGGGCCGCCAG 906
QY 427 ttctacagaagatcacaggccttctgctgtagaagattgttccattcatttagagcttctgct 486
Db 907 TTTTACCACCGGTATCGTGAGAGAGGCCCGCAGAGATTGTGCCAATTTGCTGCTGGCGAGGC 966
QY 487 tctgacagagttattgcttctctgctgaaagtctcattgaaggtttccaactctgtaagttg 546
Db 967 TCCGCGCGAGTCTATTCGTCGCGCAGAGATTCTTCAACCCGGGATTCAGGATGCCAAGAC 1026
QY 547 gctgacccaggtgcttaacccacacaaagctctccagttatttaacgttatttaccagaa 606
Db 1027 CGGGATCCCGAGGTCGAACAAAGGACACAGGAGCGCTGTGATCAACGCTGATCATTTCCGAA 1086
QY 607 ggtgctggttatacaacacactttgacaccaggttctgtgactgcttctgaaagaactctaa 666
Db 1087 GAAACTGGCAGTAACAATACTCTGGATGGGCTGACGTGCCCGCGCGCGAGGA---GGCA 1143

```

667 ttgggtgacagcgttgagtaacttaacttcaactgctgttttctgctcaacaaatagaagctaga 726
Db 1144 CCGGACCAACCCAGCCGCGAGAGTCTCTCAAGTTTTTCGCCCGCGTCTTGAAGAAAG 1203
Qy 727 ttggaagctcaactgacagtggttaacttaactgacagcaagacgttgttaacttgatgagc 786
Db 1204 ATCACTAAACACATGCGCGGTGTGAACCTCACCTTGGAGGATGTCCTGTTTCATCGAT 1263
Qy 787 atgtgtccattgacactgt-----tgtagaacttctgacgtactcaattgtctccca 840
Db 1264 CTTTGTCTGTTGACACGGTGGGCTCGACCCAGTCTTTTCCACGCGAGCTCTCTCCG 1323
Qy 841 ttctgtgactgttcaactgacagcaatgagattcaatacagactacttgcgaatttttgggt 900
Db 1324 TTTTGTCACTTGTTCAGCGGCGAGGATGAGTGGATGGCTTACGATTACTACACCCCTCGAC 1383
Qy 901 aagtactacagttacagtgctgtgtaacccattgggtccagctcaagtggtgttggttggt 960
Db 1384 AAATACTACAGCCAGCGGCGGCGAGCGCATTTGCGCGTCCCGGCGGCTCGGTTGCTC 1443
Qy 961 aacgaattgattgattgac---Lcactctccagttcaagaccacacttctactaac 1017
Db 1444 AACGAGCTGATTGCGGTATGACGGGAAATCTTCCGTTCAAGGACCCACACACAGTCAAC 1503
Qy 1018 cacactttgactctaacacagctacttccattgaagctacttctgacgtgacttc 1077
Db 1504 CACACTCTCGATGACAAACCCGGAACCTTCCCGTTGGACGCTGCTCTACGCGAGCTTT 1563
Qy 1078 tctcacacacataagtttctatttcttctgcttgggtttgttgaacagctactaac 1137
Db 1564 TCGCAGCAGACACACATGACGGGCACTTTTCCGCAATGGGCCCTGTACACGGGCACAAAG 1623
Qy 1138 ccattgtactacttctgttgaatct-----attgaagaaactgacggttactctgct 1191
Db 1624 CCGCTGTGCGAGCTCAAGATTACGCTCCGACGGGTGCGAGCGGATGGATATGCGGCA 1683
Qy 1192 tcttgactgttccattgctgctgtagagcttaagttgaatgagtgcaatgtaagagctgaa 1251
Db 1684 TCGTGGAGGTGCGCTGATGCTGCGAGCGGCGTATGTGGAGTTGCTGCGATGTGAGACGGAA 1743
Qy 1252 aag-----gaaccattggttagagtttgggttaac 1281
Db 1744 AGGACTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1803
Qy 1282 gacagagttgttccattgacaggttgggtgttgacaaagttgggttagatgtaagagagac 1341
Db 1804 GATCGGGTTGTGCGCTGATGCTGCGGTTGATGATGCGGAGGAGGAGGAGGAGGAGGAT 1863
Qy 1342 gacttcgtgaggttgttcttccgctagatctggtgttaactggaagaatgttt 1397
Db 1864 GAGTGGATTAAGGGACTCACGTTTCTCGACAGGTTGGGATTTGGGATGCTGCTT 1919

RESULT 14

US-09-163-642-1
; Sequence 1, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6221644 No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,642
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/819,825
FILING DATE: 18-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4758,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-163-642-1

Query Match 24.6%; Score 346; DB 4; Length 2200;
Best Local Similarity 57.2%; Pred. No. 3.le-93;
Matches 753; Conservative 0; Mismatches 515; Indels 48; Gaps 5;

Qy 127 gaaattctcactgtgggtacatactctcattctcttcttggctgacgaactctgt 186
Db 607 GACATTTGCCCGCACTGGGCGCAGTACTCGCCCTTCTTCTCGCTGCCGCGAGTCTCTGAA 666
Qy 187 attctccagacgttccaaaggggttagagttacttcttcttcaagtttcttagacac 246
Db 667 ATCTCGCTGCGGTGCCCAAGGGCTGCTGTCGAGTGTGTGACGCTGTGTCCCGGCAC 726
Qy 247 ggtgctagatacccaacttcttctgctgctgtaaggcgactctgttctgattgaagctatt 306
Db 727 GGAGTCTCGTATCTTACTCTCTCACAAGAGTGAAGTCTACGCGGAGTGTCTTCAAGGATC 786
Qy 307 caaagaacgtactgttccaaaggttaagtaacgtcttcttgaagacttacaactacact 366
Db 787 CAGGACATGCGACCGAGTTCAAGGCGCATTTTGGCTTCTCCGAGACTATCCCTATCAT 846
Qy 367 ttgggtgctgacgtactgactccattcgtgtgacacaaatggttaactctgtgattaaag 426
Db 847 CTCGCTGCCGATAATTTGACGGCTTTGGCGAGGAGCAGATGATGGAATCGGGCGCCAG 906
Qy 427 ttctacagaagatacaagccttggctgtagaagattgttccattcttagagcttctgt 486
Db 907 TTTCTACCCCGGTATCTGTCGAGCAGCGCCGAGAGATTGTGCCATTGTGCTGCGCGCAGGC 966
Qy 487 tctgacagagttattgtcttctgctgaaagttcattgaaggtttccaaactctgtaagttg 546
Db 967 TCCGCGCGAGTCAATGCGTCCGCGAGAGTTCTTCAACCGCGGATTCAGGATGCCAAAGAC 1026
Qy 547 gctgacccaggtgttaacccacacaaagcttctccagttatttaacgattatttccagaa 606
Db 1027 CGGGATCCCGAGTCCGAACAAGGACCGGAGGAGCCTGTGATCAACGTGATCATTTCCGAA 1086
Qy 607 ggtgctggttacacaacacttggaccacaggtttgtgttactgttctcgaagaatctgaa 666
Db 1087 GAACTGGCAGTAAACAATACTCTGGATGGCTGACGTGCCCGCGCCGAGGA---GGCA 1143
Qy 667 ttgggtgacagcttgaagctaaactcactgctgttttctgctcccaacttagagctaga 726
Db 1144 CCGGACCAACCCAGCCCGCAGAGTCTCTGCAAGTTTTCGCCCGCGCTGTCTTGAAGAAAG 1203
Qy 727 ttggaagctcacttgcaggttgttaactgactgacgaagagcttgaactgagctgagc 786

Db 1204 ATCACTAAACACATGCGGGGTGTGAACCTCACTTGCAGGATGTCCTGGTGTTCATGGAT 1263
QY 787 atgtgtccattcacactgt-----tgtagaactlctgacgtactcaattgtctcca 840
Db 1264 CTTTGCCGTTTACACAGGTGGCTCCGACCCAGTCTTTTCCAGCGGAGCTCTCTCGG 1323
QY 841 tctgtgactgttctactcacgacgaatggtattcaatagactacttgcattcttgggt 900
Db 1324 TTTTGTCACTTGTTCACGCGCCAGCATGGATGGCTTACGATTACTACTACACCTCGAC 1383
QY 901 aagctacggttacgtgtctgtaaccattgggtccagctcaaggtgtgttgggttgg 960
Db 1384 AAATACTACAGCCAGCGCGGCGGACGCAATTTGGCCCGTCCCGCGGCTCGGTTCGTC 1443
QY 961 aacgaattgattgttagattgac---tcaactctccagttcaagaccacacttctactaac 1017
Db 1444 AAGGAGCTGATTGCGGTATGACGGGAATCTCCCGTCAAGACCCACACACATCAAC 1503
QY 1018 caactttggaacttaaccagctacttcccaattgaacgtacttctgtaagcttctc 1077
Db 1504 CACACTCTCGATGACAACCCGGAATCTCCCGTTGGAGCTGTCTCTACGACACTTT 1563
QY 1078 tctcacgacacactatgttctctattttcttctgttgggtttgatacaaggttactaag 1137
Db 1564 TCCAGCAGCAACACCATGACGGGCATCTTTCCGCAATGGGCGCTGTACAACGGCACAAAG 1623
QY 1138 ccattgtactacttctgttgaatct-----attgaagaaactgacgttactctgct 1191
Db 1624 CCGCTGTGACGCTCAAGATTACGCTCCGAGCGGTGACGAGCGGATGATATGCGGCA 1683
QY 1192 tcttggaactgtccattctgctgtgtagacttacgttgaaatgtgaaagctgaa 1251
Db 1684 TCGTGGAGCGTCCCGTTCGACGAGGCGTATGTGGAGTTGCTGCGATGTGAGACGGAA 1743
QY 1252 aag-----gaaccattggttagagtttggtaac 1281
Db 1744 AGAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1803
QY 1282 gacagagttgtccattgacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
Db 1804 GATCGGTTGTCCCGCTGCATGTGTCGGGTGTGATCGATGGGAGGAGGAGGAGGAGGAG 1863
QY 1342 gactcgttgaaggttgttcttgcgtagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1397
Db 1864 GAGTGGATTAAGGAGTCACTGTTTCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1919

RESULT 15

US-09-221-654-1

; Sequence 1, Application US/09221654

; Patent No. 6054306

; GENERAL INFORMATION:

; APPLICANT: Lassen, Soren

; APPLICANT: Bech, Lisbeth

; APPLICANT: Fuglsang, Claus

; APPLICANT: Ohmann, Anders

; APPLICANT: Breinholt, Jens

; APPLICANT: Ostergaard, Peter

; TITLE OF INVENTION: Peniophora Phytase

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6054306 No. 6054306disk of No. 6054306th America, Inc.

; STREET: 405 Lexington Avenue

; CITY: New York

; STATE: NY

; COUNTRY: U.S.A.

; ZIP: 10174

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/221,654
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 08/989,358
; FILING DATE: 12-DEC-1997
; APPLICATION NUMBER: 0529/97
; FILING DATE: 07-MAY-1997
; APPLICATION NUMBER: 60/046,081
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 5101.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-221-654-1

Query Match 9.4%; Score 131.4; DB 3; Length 1320;

Best Local Similarity 56.9%; Pred. No. 1.5e-29;

Matches 268; Conservative 0; Mismatches 191; Indels 12; Gaps 1;

QY 663 tgaattgggtgacgacggtgaaagtaacttaactgctgttttgcctcccaattagac 722
Db 594 TGAAGTGGATGGTGCAGCAATCCACAAACGGTGGGGTCTTTGGCCGCAACATCACCGC 653
QY 723 tagattggaagctcacttgccaggtgttaacttgactgacgacgagctgttgaacttgat 782
Db 654 GCGATTGAACGCTGTGCGCCGAGTGCCAACTCTCAGACAGCGAGCGCTCACTCTCAT 713
QY 783 ggaatgtgtccattgcgacactgtgtgtagaactttgtgacgttacttcaattgtctccatt 842
Db 714 GGATATGTGCCCTTCGACACTCTCAG-----CTCCGGGAAGCCGACGCCCTT 761
QY 843 ctgtgactgttcaactcacgacgaatggattcaatacgaactacttgcatttttgggtaa 902
Db 762 CTGTGACCTATTTCACCGGAGGAGTATGTGCTGACGAGTACTACTATGACCTCGACAA 821
QY 903 gtactacggttacgt 962
Db 822 GTACTATGGCAGGCGCCCGGAGCGCTCTCGTCTGTCCAGGGCGGTCGGATAGTCAAA 881
QY 963 cgaattgattgttagattgactcaactctccagttcaagaccacacttctacttaaccacac 1022
Db 882 TGAGTGTGTTGACGCTTGACCGGCCAAGCCGTTTCAGACGAGAGACGAGACGACGAC 941
QY 1023 ttgtgacttaaccagctacttctccattgaacgctacttgtacgtgacttctctca 1082
Db 942 GCTCGACAGCGACCTGCAACATCCCGTGAACCTCCCGTGAACCTGACGCTTCTACGCGGACTTCGCA 1001
QY 1083 cgacaacactatggtttctatttttcttgcgtttgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1133
Db 1002 TGATAACACCATGGTGGCCATCTTTGCGGCGCTCGGGCTCTTCAACGCGCAC 1052

Search completed: October 26, 2001, 15:20:39

Job time: 595 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:10:29 ; Search time 48.78 Seconds
(without alignments)
729.264 Million cell updates/sec

Title: US-09-488-265-31

Perfect score: 2468

Sequence: 1 MGVEVLLSTATLFGSTSGT.....DFVEGLSPARSGGNWEECPA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1852	75.0	467	1 JN0889	3-phytase (EC 3.1.1)
2	1848	74.9	467	1 JN0656	3-phytase (EC 3.1.1)
3	1787	72.4	441	1 JN0482	3-phytase (EC 3.1.1)
4	405	16.4	467	1 PABYC	acid phosphatase (
5	394	16.0	467	2 S53476	acid phosphatase (
6	394	16.0	467	2 S48996	acid phosphatase (
7	391	15.8	467	1 PABYCC	acid phosphatase (
8	369	15.0	468	2 S52495	acid phosphatase h
9	360.5	14.6	479	1 JN0715	3-phytase (EC 3.1.1)
10	354.5	14.4	479	1 JN0890	acid phosphatase (
11	350	14.2	468	2 JC4285	acid phosphatase (
12	328	13.3	453	1 A25326	acid phosphatase (
13	311.5	12.6	463	2 T39929	thiamin-repressibl
14	303.5	12.3	463	2 S14119	acid phosphatase (
15	153	6.2	465	2 JE0369	histidine acid pho
16	149.5	6.1	468	2 A86233	hypothetical prote
17	142	5.8	755	2 T19118	acid phosphatase h
18	130	5.3	449	2 T15933	hypothetical prote
19	127.5	5.2	423	2 A33395	acid phosphatase (
20	118.5	4.8	421	2 S14742	acid phosphatase (
21	118	4.8	380	2 T16883	hypothetical prote
22	115.5	4.7	381	2 JH0152	acid phosphatase (
23	114.5	4.6	423	1 S06167	acid phosphatase (
24	109.5	4.4	730	2 T16455	hypothetical prote
25	107	4.3	354	2 T21241	hypothetical prote
26	106.5	4.3	602	2 A36715	exo-poly-alpha-gal
27	103	4.2	452	2 T20556	hypothetical prote
28	102.5	4.2	1225	2 T16346	hypothetical prote
29	101.5	4.1	413	2 JV0087	glucose-1-phosphat

30 101.5 4.1 413 2 B85636
31 101.5 4.1 642 2 E69144
32 101 4.1 408 2 T20893
33 100.5 4.1 630 2 S77346
34 99 4.0 386 1 JH0610
35 99 4.0 438 2 S64682
36 99 4.0 1021 2 T08601
37 98.5 4.0 4924 2 T50176
38 98 4.0 844 2 T37690
39 97.5 4.0 693 1 S61067
40 97 3.9 454 2 H82928
41 97 3.9 776 1 H69430
42 96.5 3.9 496 2 T44987
43 96.5 3.9 1433 2 A71444
44 96 3.9 721 2 T27570
45 95.5 3.9 609 1 KSPSCY

periplasmic glucos
probable formate C
hypothetical prote
hypothetical prote
acid phosphatase (C
acid phosphatase (C
hypothetical prote
probable peptide s
hypothetical prote
homocarnitate hydr
ATP synthase beta
probable formate C
aldehyde dehydroge
probable LTR retro
hypothetical prote
copper resistance

ALIGNMENTS

RESULT 1

JN0889

3-phytase (EC 3.1.3.8) A precursor - Aspergillus awamori

N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phya protein

C:Species: Aspergillus awamori

C>Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999

C:Accession: JN0889

R:Piddington, C.S.; Houston, C.S.; Paloheimo, M.; Cantrell, M.; Miettinen-Oinonen, A.

Gene 133, 55-62, 1993

A:Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-op

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0889

A:Molecule type: DNA

A:Residues: 1-467 <PID>

A:Cross-references: GB:L02421; NID:gl66518; PIDN:AAAL6898.1; PID:gl66519

A:Experimental source: strain ALK0243

A:Note: Part of the sequence, including the amino end of the mature protein, was conf

C:Comment: This enzyme catalyzes the conversion of phytate to inositol and inorganic

C:Genetics:

A:Gene: phya

A:Introns: 15/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-467/Product: 3-phytase A #status experimental <MAT>

F:27-59,105,120,207,230,339,352,376,388/Binding site: carbohydrate (Asn) (covalent) #

F:81,361/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 75.0%; Score 1852; DB 1; Length 467;

Best Local Similarity 74.3%; Pred. No. 1.le-137;

Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALPGNHSCTVDGGYOCFPEISHLMCTYSPFFSLADE 60

Db 1 MGSAVALLPLYLLAGVTSLGLAVPARNQSTCTVDQGCFTSHLMGQYAPFFSLANE 60

QY 61 SAISPDPVKGCRVTFVQVLSRHCARYPTSSASKAYSALIEAQKNATAFKGKYAFELKTYN 120

Db 61 SAISPDPVAGCRVTTAQVLSRHGARYPTESKCKYSALIEEQNNVTTFDCKYAFELKTYN 120

QY 121 YTLGADLLTPFGEOQWNSGKIFRYRYKALARKIVPFTIRASGSDRVIAAEKFIQFOSA 180

Db 121 YSLGADLLTPFGEOELVNSGKIFRYRYEYELTNIIPFTIRSSGSSRVIAAEKFIQFQST 180

QY 181 KLADPGANPHQASPVINVTPEGAGYNNLTLDHGLCTAFEESELGLDGDVEANETAVFAPPIR 240

Db 181 KLKDPRAGQGSQPKIDVVIVSEASSNNLTDPGCTCTFEDSELADTVEANFTATFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSLSPFCDLFTHDEWIQDYQLQSLG 300

Db 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSLSPFCDLFTHDEWIQDYQLQSLG 300

Db 241 ORLENDLSGVTLTDEVTYLMDCSFDTISTSTVDTKLSPPCDLTFHDEWHYDLSQSLK 300
 QY 301 KYGAGNPLGAGGVFNELIARTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGAGNPLGAGGVFNELIARTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
 QY 361 HDNTMVSIFPAGLVNKGKPLSTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK 420
 Db 361 HDNTMVSIFPAGLVNKGKPLSTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK 420
 QY 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEVEGLSFARSGGNWEECEFA 467
 RESULT 2
 JN0656
 3-phytase (EC 3.1.3.8) A precursor - Aspergillus niger
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phya protein
 C:Species: Aspergillus niger
 C:Date: 03-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 11-Jun-1999
 C:Accession: JN0656; S28456
 R:van Hartingsveldt, W.; van Zeijl, C.M.J.; Harteveld, G.M.; Gouka, R.J.; Suykerbuyk, M.
 A.M.J.J.
 Gene 127, 87-94, 1993
 A:Title: Cloning, characterization and overexpression of the phytase-encoding gene (phyA)
 A:Reference number: JN0656; MUID:93252284
 A:Accession: JN0656
 A:Molecule type: DNA
 A:Residues: 1-467 <NAN>
 A:Cross-references: GB:216414; MID:q2392; PIDN:CAA78904.1; PID:q2393
 A:Experimental source: strain NRRL3135
 A:Note: parts of the sequence, including the amino end of the mature protein, were confirmed
 C:Comment: This enzyme catalyzes the hydrolysis of phytic acid into myo-inositol and inositol
 C:Genetics:
 A:Gene: phya
 A:Introns: 15/2
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; phosph
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-467/Product: 3-phytase A #status experimental <MAT>
 F:27-59,105,120,207,339,352,376,386/Binding site: carbohydurate (Asn) (covalent) #sta
 F:81,361/Active site: Arg, His #status predicted
 F:82/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 74.9%; Score 1848; DB 1; Length 467;
 Best Local Similarity 74.3%; Pred. No. 2.2e-137;
 Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLLSTATLFGSTGALGPRGNSHSDTVDGVCYQCFPEISHLWGTYSPPFSLADE 60
 Db 1 MGVSALLLPLLYLLSGVTSGLAVPASRNSQSDTVDGVCYQCFSETSHLWQYAPFSLANE 60
 QY 61 SATSPDVPKGRVTFVQVLSRGARYPTSSASKAYSALIEATQKNATAPKGYAFLKTYN 120
 Db 61 SVISPEVPAGCRVTFQVLSRGARYPTDSKGYKYSALIEEQQNTATFDGKYAFLKTYN 120
 QY 121 YTLGADLLPFGQMVNSGIKRYKALARKIVPIRASGSDRVIASAEKFIQGFOSA 180
 Db 121 YSLGADLLPFGQELVNSGIKRYKESLRTNIVPIRSSSSSRVIAKGFIEGFQST 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 KLADPRAQPGQSPKIDVISEASSNNTLDPCTCTVPEDSELADVEANFTAVPPIR 240
 QY 241 ARLEAHLPGVNLTDDEVNLMCMCPDPTVARTSDATQLSPFCDLTFHDEWHYDLSQSLK 300
 Db 241 ORLENDLSGVTLTDEVTYLMDCSFDTISTSTVDTKLSPPCDLTFHDEWHYDLSQSLK 300
 QY 301 KYGAGNPLGAGGVFNELIARTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 KYGAGNPLGAGGVFNELIARTHSPVDHTSTNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNTMVSIFPAGLVNKGKPLSTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK 420
 Db 361 HDNTMVSIFPAGLVNKGKPLSTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK 420
 QY 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGVVDKLGCRKRDDEVEGLSFARSGGNWEECEFA 467
 RESULT 3
 JN0482
 3-phytase (EC 3.1.3.8) A - Aspergillus ficum
 N:Alternate names: myo-inositol hexakisphosphate phosphohydrolase; phya protein
 C:Species: Aspergillus ficum
 C:Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 07-May-1999
 C:Accession: JN0482; PN0023
 R:Ullah, A.H.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 192, 747-753, 1993
 A:Title: Aspergillus ficum phytase: Complete primary structure elucidation by chemi
 A:Reference number: JN0482; MUID:93249451
 A:Accession: JN0482
 A:Molecule type: protein
 A:Residues: 1-441 <ULL>
 A:Note: authors state that the 9 Asn followed by Thr or Ser after an intervening resi
 R:Ullah, A.H.J.; Cummins, B.J.; Dischinger Jr., H.C.
 Biochem. Biophys. Res. Commun. 178, 45-53, 1991
 A:Title: Cyclohexanediol modification of arginine at the active site of Aspergillus
 A:Reference number: PN0023; MUID:91298982
 A:Accession: PN0023
 A:Molecule type: protein
 A:Residues: 48-70 <UL2>
 A:Comment: This enzyme catalyzes the hydrolysis of inorganic orthophosphate from phyt
 C:Superfamily: yeast acid phosphatase
 C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph
 F:4,36,82,97,184,207,316,329,353,365/Binding site: carbohydurate (Asn) (covalent) #sta
 F:58,338/Active site: Arg, His #status predicted
 F:59/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 72.4%; Score 1787; DB 1; Length 441;
 Best Local Similarity 75.8%; Pred. No. 1.3e-132;
 Matches 332; Conservative 40; Mismatches 66; Indels 0; Gaps 0;

QY 27 NNSHCDVVDGVCYQCFPEISHLWGTYSPPFSLADESAISPDVPKGRVTFVQVLSRGARY 86
 Db 4 NNSHCDVVDGVCYQCFSETSHLWQYAPFSLANESVISPEVPAGCRVTFQVLSRGARY 63
 QY 87 PTSASAKAYSALIEATQKNATAPKGYAFLKTYNLTGADLLPFGQMVNSGIKRYR 146
 Db 64 PTDSKGRKYSALIEEQQNTATFDGKYAFLKTYNLSLADLLPFGQELVNSGIKRYR 123
 QY 147 YKALARKIVPIRASGSDRVIASAEKFIQGFQSAKLADPGANPHQASPVINVIPEGAGY 206
 Db 124 YESLRTNIVPIRSSSSSRVIAKGFIEGFQSTKLDPRAPQCSQSPKIDVISEASS 183
 QY 207 NNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRLEAHLPGVNLTDDEVNLMCMCP 266
 Db 184 NNTLDPGTCTVFESELDATVEANFTATVPFSRQRLENDLSGVTLTDEVTYLMDCSF 243
 QY 267 DTAVARTSDATQLSPFCDLTFHDEWHYDLSQSLKYYGYGAGNPLGPAQGVFNELIAR 326
 Db 244 DTISTSTVDTKLSPPCDLTFHDEWHYDLSQSLKYYGYGAGNPLGPTQGVGVANELIAR 303
 QY 327 LTHSPVDHTSTNHTLDSNPATFPLNATLYADFSHDNTMVSIFPAGLVNKGKPLSTTSV 386
 Db 304 LTHSPVHDDTSSNHTLDSNPATFPLNATLYADFSHDNGIISILFALGLYNGTKPLSTTV 363
 QY 387 ESIEETDGYASWTVPFAARAYVEMMOCEAEKPELVRLVNDVRVPLHGGVVDKLGCRK 446
 Db 364 ENITQDGFSSAWTVPFAARAYVEMMOCEAEKPELVRLVNDVRVPLHGGVVDKLGCRK 423
 QY 447 DDFVEGLSFARSGGNWEE 464

;Residues: 1-2,'Y',4-43,'T',45-51<BER>

Submitted to the EMBL Data Library, February 1994
;Description: Sequencing of chromosome I of *Saccharomyces cerevisiae*: analysis of th

A:Reference number: S53458
 A:Accession: S53476
 A:Molecule type: DNA
 A:Residues: 1-467 <BUS>
 A:Cross-references: EMBL:L28920; NID:g1616966; PIDN:AAC09508.1; PID:g456155; MIPS:YAR071
 R:Chen, J. Y.; Gong, Y.; Ao, S. Z.
 Acta Biochim. Biophys. Sin. 21, 437-444, 1989
 A:Title: The primary structure of acid phosphatase gene PHO11 in *S. cerevisiae* and comparison with other members of the acid phosphatase gene family
 A:Reference number: JC1018
 A:Accession: JC1018
 A:Molecule type: DNA
 A:Residues: 1-16, 'L', '18-149, 'H', '151-353, 'Q', '355-422, 'G', '424-467 <CHE>
 A:Note: this paper is in Chinese, with an English abstract
 C:Genetics:
 A:Gene: SGD:PHO11
 A:Cross-references: SGD:S0000094; MIPS:YAR071w
 A:Map position: 1R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-467/Product: acid phosphatase #status predicted <MA>
 F:74/Active site: Arg #status predicted
 F:75/Active site: His (phosphohistidine intermediate) #status predicted
 F:97,162,192,250,315,356,439,445,461/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 16.0%; Score 394; DB 2; Length 467;
 Best Local Similarity 25.7%; Pred. No. 3.7e-23;
 Matches 119; Conservative 65; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCPPEISHLWGTYSPPFSLADESAISDPVPGKGRVTFVQVLSRHG 83
 DB 21 PLGKLSDDIKGTQTEIFPFL----GGSPYYSFPGDYGISRDLPESCEMKQVQVGRHG 76

QY 84 ARPTSSASKAYSAALIEAQKNATAPKGYAFLK-----TNYTL-----GADDL 128
 DB 77 ERYPTVSKAKSIMTWTYKLSNVTGQFSGALSFNDYEFFIRDTKMLEMETTLANSVNVL 136

QY 129 TPF-GEQMVNSGIFRYRYKALARKIVPF-IRASGSDRVIASAEKFIETGFSQAKLADPG 186
 DB 137 NPYTGEMNAKRIARDELAQYGVNENQTSFAVTSNSNRCHDTAQYFDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241
 DB 191 -----FNISLQPISEASAGANTLSAHSACPWDD-DVNDLILKKYDKYLSGLIAK 240

QY 242 RLEAHLPGVNLTDVNVNMDMCPDFTVARTSDATQLSPFCDLTHDEWIQYDYLSLQSGK 301
 DB 241 RLKNGKGLNLTSSDANTFFAWCAYEINAR-----GYSDICNIFTKDELVRFSYGODLET 295

QY 302 YGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSH 361
 DB 296 YYQTGPGYDVRSVGVANLFNASVKLLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFGALNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEKEPL 421
 DB 338 DTDILNLTITIGIIDDKNLTAEHVPEMENT--FHRSWYVPOGARVYTEKFCQ--SNDTY 393

QY 422 VRVLVNDRVVPLHCGGVDKLGCRKDDF-----VEGLSFAR 457
 DB 394 VRYVINDAVVPIETCTGPGFSCFNDYFAEKRKRVAGTDFLK 436

RESULT 6
 S48996
 acid phosphatase (EC 3.1.3.2) PHO12 - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein YHR215w
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 22-Jun-1999
 C:Accession: S48996; S59659
 R:Macri, C.
 submitted to the EMBL Data Library, February 1994
 A:Description: The sequence of *S. cerevisiae* cosmid 9177.

A:Reference number: S46671
 A:Accession: S48996
 A:Molecule type: DNA
 A:Residues: 1-467 <MAC>
 A:Cross-references: EMBL:U00029; NID:g551322; PIDN:AAB69729.1; PID:g458917; MIPS:YHR2
 R:Yu, L.
 submitted to the EMBL Data Library, January 1995
 A:Reference number: S59658
 A:Accession: S59659
 A:Molecule type: DNA
 A:Residues: 1-16, 'L', '18-81, 'AR', '84-149, 'H', '151-467 <XUL>
 A:Cross-references: EMBL:U19789; NID:g847754; PIDN:AAA73479.1; PID:g847755
 C:Genetics:
 A:Gene: SGD:PHO12
 A:Cross-references: SGD:S0001258; MIPS:YHR215w
 A:Map position: 8R
 C:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 16.0%; Score 394; DB 2; Length 467;
 Best Local Similarity 25.7%; Pred. No. 3.7e-23;
 Matches 119; Conservative 65; Mismatches 203; Indels 76; Gaps 14;

QY 24 PRGNHSCDVTGQYQCPPEISHLWGTYSPPFSLADESAISDPVPGKGRVTFVQVLSRHG 83
 DB 21 PLGKLSDDIKGTQTEIFPFL----GGSPYYSFPGDYGISRDLPESCEMKQVQVGRHG 76

QY 84 ARPTSSASKAYSAALIEAQKNATAPKGYAFLK-----TNYTL-----GADDL 128
 DB 77 ERYPTVSKAKSIMTWTYKLSNVTGQFSGALSFNDYEFFIRDTKMLEMETTLANSVNVL 136

QY 129 TPF-GEQMVNSGIFRYRYKALARKIVPF-IRASGSDRVIASAEKFIETGFSQAKLADPG 186
 DB 137 NPYTGEMNAKRIARDELAQYGVNENQTSFAVTSNSNRCHDTAQYFDG-----LGDK- 190

QY 187 ANPHQASPVINII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241
 DB 191 -----FNISLQPISEASAGANTLSAHSACPWDD-DVNDLILKKYDKYLSGLIAK 240

QY 242 RLEAHLPGVNLTDVNVNMDMCPDFTVARTSDATQLSPFCDLTHDEWIQYDYLSLQSGK 301
 DB 241 RLKNGKGLNLTSSDANTFFAWCAYEINAR-----GYSDICNIFTKDELVRFSYGODLET 295

QY 302 YGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSH 361
 DB 296 YYQTGPGYDVRSVGVANLFNASVKLLKESEVQDQ-----KWLSFTH 337

QY 362 DNTMVSIFGALNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEKEPL 421
 DB 338 DTDILNLTITIGIIDDKNLTAEHVPEMENT--FHRSWYVPOGARVYTEKFCQ--SNDTY 393

QY 422 VRVLVNDRVVPLHCGGVDKLGCRKDDF-----VEGLSFAR 457
 DB 394 VRYVINDAVVPIETCTGPGFSCFNDYFAEKRKRVAGTDFLK 436

RESULT 7
 PABYCC
 acid phosphatase (EC 3.1.3.2) precursor, constitutive - yeast (*Saccharomyces cerevisi*
 N:Alternate names: acid phosphatase PHO3; protein YBR0813; protein YBR092c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 30-Sep-1991 #sequence_revision 09-Sep-1994 #text_change 05-Nov-1999
 C:Accession: S48259; S05794; A25241; S44674
 R:Mannhaupt, G.; Stucka, R.; Ehnl, S.; Vetter, I.; Feldmann, H.
 Yeast 10, 1363-1381, 1994
 A:Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
 A:Reference number: S48255; MUID:95208357
 A:Accession: S48259
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-467 <MAN>
 A:Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55597.1; PID:g476050

Db 423 DYAEKRVAGTDFLK 436

```

;Date: 14-Jul-1994 #sequence_revision 19-Oct-1995

```

3-phytase (EC 3.1.3.8) B precursor - *Aspergillus ficuum*
 N/Alternate names: pH 2.5-Optimum acid phosphatase
 C/Species: *Aspergillus ficuum*
 C/Date: 14-Jul-1994 #sequence_revision 19-Oct-1995 #text

A>Title: The cloning and sequencing of the genes encoding phytase (phy) and pH 2.5-o-

A:Reference number: JN0889; MUID:94040796

A:Accession: JN0890

A:Molecule type: DNA

A:Residues: 1-479 <PID>

A:Cross-references: GB:I02420; NID:g166481; PIDN:AAAI6897.1; PID:g166482

A:Experimental source: strain ALK0243

C:Comment: The highly similar enzyme from A. ficuum has been shown to have 3-phytase

C:Genetics:

A:Gene: aph

A:Introns: 261/1; 300/2; 335/2

C:Superfamily: yeast acid phosphatase

C:Keywords: extracellular protein; glycoprotein; phosphohistidine; phosphoprotein; ph

F:1-Domain: signal sequence #status predicted <SIG>

F:20-479/Product: 3-phytase #status predicted <MAT>

F:81,337/Active site: Arg, His #status predicted

F:82/Active site: His (phosphohistidine intermediate) #status predicted

F:106,191,227,250,315,340,425,442,458/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 14.4%; Score 354.5; DB 1; Length 479;
Best Local Similarity 27.9%; Pred.No. 4.8e+20;
Matches 117; Conservative 53; Mismatches 159; Indels 91; Gaps 15;

Qy 49 GTYSPFFSLADESAISPDVPKGKRVTEVQLSRHGARYPTSSASK-AYSALIEIAIQRNAT 107
||| ||| : : : | : : : ||| ||| : : : ||| : :
Db 53 GPYSERSVS---GIARDPPTSCEVDQVIMVKRHGERYPSPSAGKDIEEALKVYSINTT 108
||| ||| : : : | : : : ||| ||| : : : ||| : :

Qy 108 AFKKYAFLKYNTYL-----GADDTL-PF-GEOQMNSGIKFRRYKAL--ARKIVPF 157
:|| ||| : : : | : : : ||| ||| : : : ||| : :
Db 109 EYKGDLAFLDWTYYVNECYNAETTSFGYAGLLDAYNHGNDYKARYGHILWNETGVVPF 168
||| ||| : : : | : : : ||| ||| : : : ||| : :

Qy 158 IRASGSDRVTASAEEKFEQGSAKLDPGANPHQASPVINVIPEGAGYNLTLDHGLCTA 217
:|| ||| ||| ||| ||| ||| :
Db 169 F-SGGYGRVLETAKRKEGF-----FCYNSTNAALNII 201
||| ||| ||| ||| ||| ||| :

Qy 218 FEESLG-----DDVEANFTAVFAPPIR----ARLEAHLPGVLNLTDEDVNLMD 262
: : : | : : : | : : : ||| : : : ||| ||| ||| |||
Db 202 SESEVMGADSLTPCSTDNDQTTCDNLTQLPQFKAAARLN SQNPGNLTASDVYNLMV 261
||| ||| : : : | : : : ||| ||| : : : ||| : :

Qy 263 MCPEDTVARTSDATQLSPFCDLFTHDEWIQDYDLQSIGKYGCGAGNPLGPACOGVGFWNE 322
: : : | : : : ||| : : : ||| ||| : : : ||| : :
Db 262 MASPELNAR-----PFSNWINAFTQDEWSFGYVEDLYNYCCAGPGDKNNMAAGVAYANA 316
||| ||| : : : | : : : ||| ||| : : : ||| : :

Qy 323 LIARLTHSPVDHTSTNHTLDSNPATEPLNATLVADFESHDMTMSIFFALGLY--NGTKP 380
: : : | : : : ||| : : : ||| ||| : : : ||| : :
Db 317 SLITLNOCPKE-----AGSLFNFAHDNTIIPILAAALGVLI PNEDLP 358
||| ||| : : : | : : : ||| ||| : : : ||| : :

Qy 381 LSTTVSVESIETDGYSASVTVPFAARAYVENMOCEA-----EKEPLVRVLVNDRVVP LHGC 436
: : : | : : : ||| : : : ||| ||| : : : ||| : :
Db 359 L-----DRVAFGPNPYSGINIVPMGGHILTIERLSQCATALSDEGTYVRLVLEAVLPFND C 413
||| ||| : : : | : : : ||| ||| : : : ||| : :

RESULT 11

JC4285

N:Altophatase (EC 3.1.3.2) precursor - yeast (*Pichia pastoris*)

N:Alternate names: orthophosphoric monoester phosphohydrolase (acid optimum); Ph

C:Species: Pichia pastoris

C>Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 29-Oct-1999

C:Accession: JC4285

R:Payne, W.E.; Gannon, P.M.; Kaiser, C.A.

Gene 163, 19-26, 1995

A>Title: An inducible acid phosphatase from the yeast *Pichia pastoris*: Characterizat

A:Reference number: JC4285; MUID:96001238

A:Accession: JC4285

A:Molecule type: DNA

A:Residues: 1-468 <PAY>

A:Cross-references: GB:U28658; NID:g981955; PIDN:AAA85503.1; PID:g981956

A:Experimental source: GS115

C:Genetics:

A:Gene: pho1

C:Superfamily: yeast acid phosphatase

C:Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hyd

RESULT 12
 A25326
 acid phosphatase (EC 3.1.3.2) precursor [similarity] - fission yeast (*Schizosaccharomyces*
C. species: Schizosaccharomyces pombe
 C:date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C/accession: A25326; T50405
 R; Elliott, S.; Chang, C.; Schweingruber, M.E.; Schaller, J.; Rickli, E.E.; Carbon, J.
 J. Biol. Chem. 261, 2936-2941, 1986
 A; title: Isolation and characterization of the structural gene for secreted acid phosphatase
 A; reference number: A25326; MUID:86140050
 A; accession: A25326
 A; molecule type: DNA
 A; residues: 1-453 <ELL>
 A; cross-references: GB:M11857; NID:gl73422; PIDN:AAA35321.1; PTD:gl73423
 R; Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 2000
 A; reference number: Z25068
 A; accession: T50405
 A; molecule type: DNA
 A; residues: 1-453 <RIE>
 A; cross-references: EMBL:AL137099; PIDN:CAR68657.1; GSPDB:GN00067; SPDB:SPBP4G3.02
 A; experimental source: strain 97Zh(-); clone p1 p4G3
 C; genetics:
 A; gene: phol; SPDB:SPBP4G3.02
 A; map position: 2

```

QY      46  HLKGTYS-----PPFSLADESAISDPVPGKRVITFVQLSRHGARYPTS--SASKAYSA--97
      35  HL-GTSPVHEPFDGLSDA-----EPETCEIQQVHLLGRGSRNPTGVDTATDVSSOY  88
QY      98  -----LIEATQKNATAKGYAFLKTYNTPL--GADDTLPFGQQMWSNGIKFYRR  146
      89  LNNFQKLLNGSPVNFSPENPLCFIKQWTPVIDENADQLSSRCRLEFLDGLGLOYR  148
QY     147  YKALARKIVPFIIRASGSDRVIAAEKFIEGFOSAKIADPAGNPHOASPVYNVIIPFG--A  204

```

Db 149 YKLEDSYVDINTAEQKRVESAKWTFYGLFGDKWYE-----KTNFILLISEGKAA 199
 QY 205 GYNNTLDHGLCTAFBESEL----GDDVEANETAVFAPPFIRARLEAHL-PGVNLTDEDVYN 259
 Db 200 GANSLSMYNACPVFKDNNPHKKNATDAHAHVNFIEPIVNRKAKYFDSYKLTINDVRS 259
 QY 260 LDMCPCFDVARTSDATQSPFCDLFTDHEWIOYDYLQSLGKYGYGAGNPLGPAQGVGF 319
 Db 260 LFVICEYEIAIKDH-----SDFCSIFTPSEFLNFYDSDLDQAYGGGPVSEWASTLGAY 314
 QY 320 VNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFAALGLYNGTK 379
 Db 315 INNLADSL-----RNVTNPDEFDRK-----VFLAFTHDSNLIIPVEAALGFPPDIT 358
 QY 380 PLSTTSVESIEETDGYASWTVPFAARAYVENMOQAEKEPLRVLVNDRVVPLHGCQVD 439
 Db 359 PQNPLPTDKNIITYSOKTSSFPFAGNLITELFFC-SDSKYYVVRHLVNOQVYPLIDCGYG 417
 QY 440 KLGR-----CKRDDFVEGLSFARSGGN 461
 Db 418 PSGTSDGLCELQAOYLNSPIRANSTSN 443

RESULT 14

SL1419
 A:Title: acid phosphatase (EC 3.1.3.2) - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
 C:Accession: SL1419; T40455
 R:Yang, J.; Schweingruber, M.E.
 Curr. Genet. 18, 269-272, 1990
 A:Title: The structural gene coding for thiamin-repressible acid phosphatase in Schizosaccharomyces pombe
 A:Reference number: SL1419; MUID:91064763
 A:Accession: SL1419

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <YAN>
 A:Cross-references: GB:X56939; NID:g5006; PIDN:CAA0258.1; PID:g5007
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z21931
 A:Accession: T40455
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-463 <LYN>
 A:Cross-references: EMBL:AL034382; PIDN:CAA22278.1; GSPDB:GN00067; SPDB:SPBC428.03c
 A:Experimental source: strain 972h; cosmid c428
 C:Genetics:
 A:Gene: SPBC428.03c
 A:Map position: 2
 A:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 12.3%; Score 303.5; DB 2; Length 463;
 Best Local Similarity 26.7%; Pred. No. 4.7e-16;
 Matches 113; Conservative 63; Mismatches 184; Indels 63; Gaps 16;

QY 46 HLMGTYS-----PFFSLADESAISPDVPKGRVTFVQVLSRHGARYPT--SSASKAYSA-- 97
 Db 35 HL-GTISVYHEPYF-----NGPTTSPFESCAIKQVHLLQHRGSRNFTGDDTATDVSSAQY 88
 QY 98 -----LIEATOKNATAFKYAFILKTYNYTL--GADDLTFPGEQOOWNSGKIFYRR 146
 Db 89 IDIFQNLKNGSIPVNFSPENFLYFVKHTVPVKEANADQLSSSGRIELFDLGRGVFR 148
 QY 147 YKALARKIVPFIIRASGSDRVIASEAIEFGFSQAKLADPAPNPHQASPVNVIPIE--GA 204
 Db 149 YYELFDYDVIDYNTAAQERVVDSAEWFSYGM-----FGDDMQNKTNF-----IVLPEDDSA 199
 QY 205 GYNNTLDHGLCTAFBESELGDD--VEANFTA---VFAPPFIRARLEAHL-PGVNLTDEDVYN 259

Db 200 GANSLAMYSCPVEYEDNNIDNTTEAAHTSWRNVFLKPIANRLNKYFDSGYNLTVSDVRS 259
 QY 260 LDMCPCFDVARTSDATQSPFCDLFTDHEWIOYDYLQSLGKYGYGAGNPLGPAQGVGF 319
 Db 260 LYXICVYEIALRDN-----SDFCSIFTPSEFLNFYDSDLDQAYGGGPVSEWASTLGAY 314
 QY 320 VNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFFAALGLYNGTK 379
 Db 315 VNNLANNL-----RKGYNNASDRK-----VFLAFTHDSQIIPVEAALGFPPDIT 358
 QY 380 PLSTTSVESIEETDGYASWTVPFAARAYVENMOQAEKEPLRVLVNDRVVPLHGCQVD 439
 Db 359 PEHPLPTDKNIITYSOKTSSFPFAGNLITELFFC-SDSKYYVVRHLVNOQVYPLIDCGYG 417
 QY 440 KLGR 442
 Db 418 PSG 420

RESULT 15

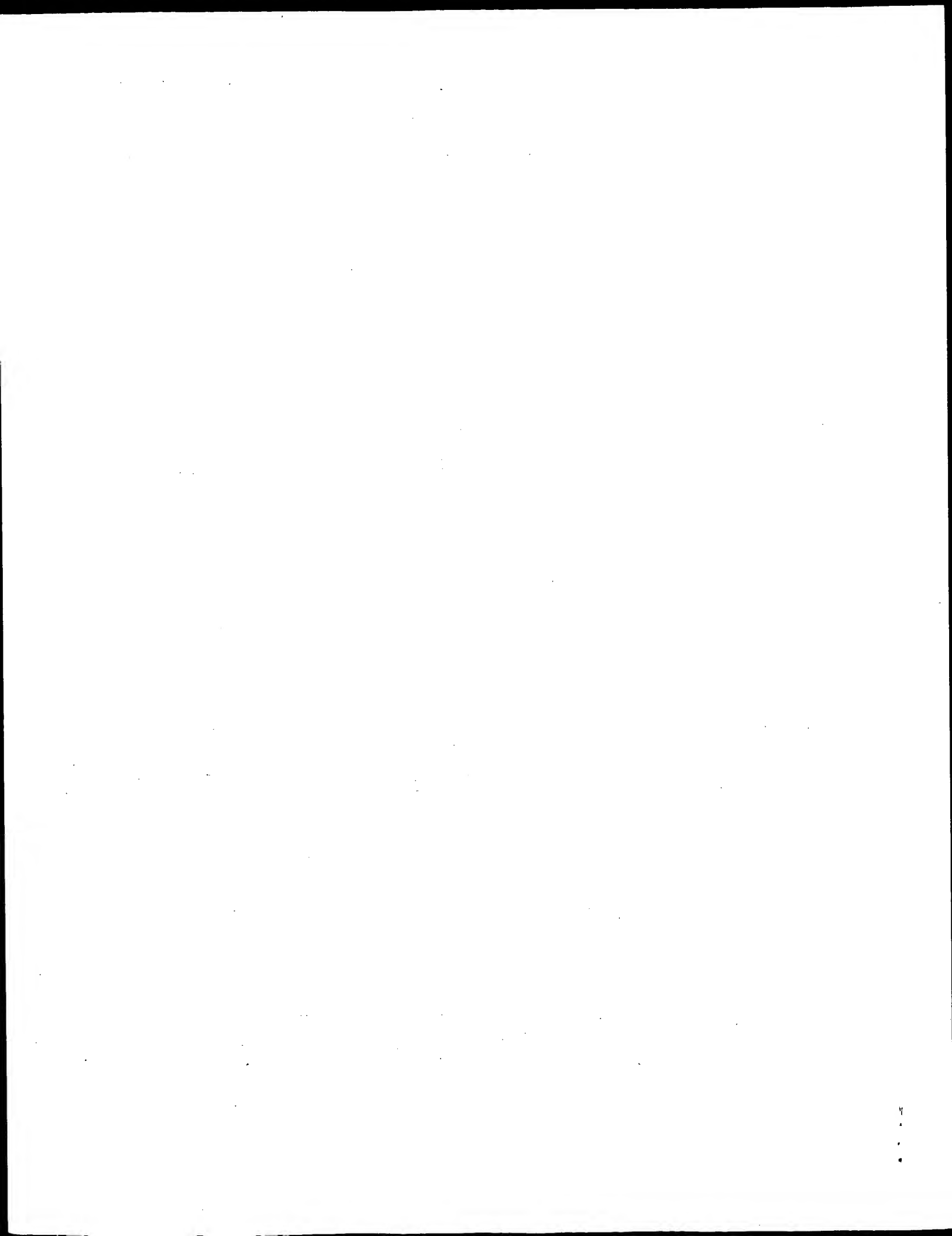
JE0369
 A:Title: histidine acid phosphatase (EC 3.1.3.-) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
 C:Accession: JE0369
 R:Mullaney, E.J.; Ullah, A.H.J.
 Biochem. Biophys. Res. Commun. 251, 252-255, 1998
 A:Title: Identification of a histidine acid phosphatase (phyA)-like gene in Arabidopsis

A:Reference number: JE0369; MUID:99009256
 A:Accession: JE0369
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-465 <MUL>
 A:Superfamily: yeast acid phosphatase
 C:Keywords: phosphoric monoester hydrolase

Query Match 6.2%; Score 153; DB 2; Length 465;
 Best Local Similarity 22.5%; Pred. No. 0.00031;
 Matches 102; Conservative 61; Mismatches 199; Indels 92; Gaps 21;

QY 35 DGYQCFPETSHLWGT---YSPFSLADESAISPDVPKGRVTFVQVLSRHGARYPTSSA 91
 Db 16 DGGF---DHRHLSTVTRYSTSKDVTQNLIEGSNVPSSECTPIHLNLVARHGTRSPTKR 71
 QY 92 SKAYSAL---LEAIQKNATAFK-----GKYAFILKTYNYTLGADDLTFPGEQOOWN 138
 Db 72 LREMESLAGRFKELVROAEARKLPDSKIPGWLQWK--SPWEGKVGKGLIRIGEDELQY 129
 QY 139 SGIKFYRYKALARK----IVPFIIRASGSDRVITASAEKFIIEGFSQAK-LADPGANPHQAS 193
 Db 130 LGIRVREFFSLFEEDYHPDVYTTIRATQIPRASASAVAFGMLFSEKGNLGPGRNRAFA- 188
 QY 194 PVINVIPEGAGYNNTLDHGLCTAFBESELGDDVEANFTAVFAPPFIRARLEAHLPGVNL 253
 Db 189 -----VTSNRSADTKLRFEECC-----QNY-----KSYRAKEPAVDKL 223
 QY 254 DEDVNLMDMC---PFDIVARTSDATQSPFCDLFTDHEWIOYDYLQSLGKYGYGAGNP 310
 Db 224 KEPVNLKITASVAKRYDLKFTKQDISLWFLCKQVALLW-----TDDLEVLKKGYSN 278
 QY 311 LGPAGGVGVNLIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMWSIFF 370
 Db 279 LNYKMGVPLLEDVL---HSMEEAIKAREEKLK--PGSYE-KARL--RFAHAETIVPPFS 329
 QY 371 A-LGLYNGTKPLSTTSVESIE-----ETDGYASWTVPFAARAYVENMOQAE 417
 Db 330 CULGLF-----LDGEFEKIQKEKPLPPOPKTRDFRGSTMAPFGNNLLVLYSCP 384
 QY 418 KEP--LVRLVNDRVVPLHGCQVDKLGKCRDDF 449
 Db 385 SSPKYFVQVHLNHEHPIAVPGC--DGKDFCPLDEF 416

Search completed: October 26, 2001, 15:10:30
Job time: 6336 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:18:04 ; Search time 29.88 Seconds

(without alignments)
535.385 Million cell updates/sec

Title: US-09-488-265-31

Perfect score: 2468.

Sequence: 1 MGVFVLLSTATLFGTSGT.....DFVEGLSFARSGGNWEECA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	1852	75.0	467	PHYA_ASPAW	P34753 aspergillus
2	1848	74.9	467	PHYA_ASPNG	P34752 aspergillus
3	1804.5	73.1	463	PHYB_EMENI	O00093 emericeella
4	405	16.4	467	PPA5_YEAST	P00635 saccharomyc
5	394	16.0	467	PPAB_YEAST	P35842 saccharomyc
6	394	16.0	467	PPAC_YEAST	P38693 saccharomyc
7	391	15.8	467	PPA3_YEAST	P24031 saccharomyc
8	369	15.0	468	PPAD_YEAST	P52290 saccharomyc
9	362.5	14.7	479	PHYB_ASPNG	P34754 aspergillus
10	354.5	14.4	479	PHYB_ASPAW	P34755 aspergillus
11	350	14.2	468	PPAL_PICPA	P52291 pichia past
12	337.5	13.7	469	PPA5_KLUJLA	P52289 kluyveromyc
13	328	13.3	463	PPA1_SCHPO	P08091 schizosacch
14	303.5	12.3	463	PPA2_SCHPO	O01682 schizosacch
15	142	5.8	755	PPAX_CAEEL	Q09549 caenorhabdi
16	127.5	5.2	423	PPAL_RAT	P20611 rattus norv
17	118.5	4.8	421	PPAL_MOUSE	P24638 mus musculus
18	115.5	4.7	381	PPAP_RAT	P20646 rattus norv
19	114.5	4.6	423	PPAL_HUMAN	P11117 homo sapien
20	106.5	4.3	602	PEHX_ERWCH	P15922 escherichia
21	101.5	4.1	413	AGP_ECOLI	P19926 escherichia
22	99	4.0	386	PPAP_HUMAN	P15309 homo sapien
23	98	4.0	844	YDZ5_SCHPO	O13712 schizosacch
24	97.5	4.0	693	LYS4_YEAST	P43367 saccharomyc
25	97	3.9	411	PPAY_CAEEL	Q10944 caenorhabdi
26	96.5	3.9	845	CC47_YEAST	P38132 saccharomyc
27	96	3.9	721	NTCA_PSEEM	Q23316 caenorhabdi
28	95.5	3.9	609	COPA_PSEEM	P12374 pseudomonas
29	94.5	3.8	526	CLOS_CLOHI	P09870 clostridium
30	94	3.8	467	ATPX_BACFI	P25075 bacillus fi
31	94	3.8	938	TRGI_ECOLI	P33790 escherichia
32	93.5	3.8	2485	PTND_HUMAN	Q12923 homo sapien
33	93	3.8	428	GSA_STAAU	O34092 staphylococ

34	93	3.8	777	1	ISOA_FLASP	O32611 flavobacter
35	92.5	3.7	433	1	DHOM_SYNY3	P52986 synchocyst
36	92.5	3.7	760	1	MCW7_SCHPO	O75001 schizosacch
37	92	3.7	3587	1	TYCB_BACBR	O30408 b tyrocidin
38	91.5	3.7	351	1	PYRD_HELPY	O25655 helicobacte
39	91.5	3.7	586	1	VP40_ILTVT	P23984 infectious
40	91.5	3.7	680	1	NCPR_CANMA	P50126 candida mal
41	91	3.7	417	1	AGP_PRORE	O52309 providencia
42	90.5	3.7	646	1	IRGI_MOUSE	P54987 mus musculu
43	90	3.6	626	1	BGAL_LEULA	Q02603 leuconostoc
44	90	3.6	1527	1	MRP3_HUMAN	O15438 homo sapien
45	89.5	3.6	492	1	CRTI_ERWHE	P22871 erwinia her

ALIGNMENTS

RESULT	1
PHYA_ASPAW	
ID	PHYA_ASPAW
AC	P34753
DT	01-FEB-1994 (Rel. 28, Created)
DT	01-FEB-1994 (Rel. 28, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	3-PHYTASE A PRECURSOR (EC 3.1.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE	3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE	PHOSPHOHYDROLASE A).
GN	PHYA OR PHY.
OS	Aspergillus awamori.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC	Euriales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX	NCBI_TaxID=105351;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ALK0243;
RX	MEDLINE=94040796; PubMed=8224894;
RA	Piddington C.S., Houston C.S., Paloheimo M.T., Cantrell M.A.,
RA	Miettinen-Oinonen A., Nevalainen H., Rambossek J.A.;
RT	"The cloning and sequencing of the genes encoding phytase (phy) and
RT	pH 2.5-optimum acid phosphatase (aph) from Aspergillus niger var.
RT	awamori.";
RL	Gene 133:55-62(1993).
CC	-!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC	FROM PHYTATE.
CC	-!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC	D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC	-!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; L02421; AAA16898.1; "
DR	PIR; JN0889; JN0889.
DR	HSSP; P34752; LIHP.
DR	InterPro; IPR000560; "
DR	Pfam; PF00328; acid.phosphat; 1.
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
DR	Hydrolase; Glycoprotein; Signal.
FT	SIGNAL 1 23
FT	POTENTIAL.
FT	CHAIN 24 467
FT	3-PHYTASE A.
FT	ACT_SITE 82 82
FT	NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT	ACT_SITE 361 361
FT	PROTON DONOR (BY SIMILARITY).
FT	DISULFID 31 40
FT	BY SIMILARITY.
FT	DISULFID 71 414
FT	BY SIMILARITY.
FT	DISULFID 215 465
FT	BY SIMILARITY.
FT	DISULFID 264 282
FT	BY SIMILARITY.
FT	DISULFID 436 444
FT	BY SIMILARITY.


```

FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 51075 MW; 118928A5D7EC661 CRC64;

Query Match 75.08; Score 1852; DB 1; Length 467;
Best Local Similarity 74.38; Pred. No. 2.1e-140;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGCVFVLLSIATLFGSGTALGPRGNHSCDVTGQYQCFPEISHLWGYSPFFSLADE 60
DB 1 MGVSALLPLLYLAGVTSGLVAPASRNQSTCDTVDQGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SAISPDVPGKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGKVAFLKTYN 120
DB 61 SAISPDVPGKGRVTFVQVLSRHGARYPTESKGYKYSALIEIQNVTFDGKVAFLKTYN 120

QY 121 YTLGADDLTPFGEOMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTIEGFOSA 180
DB 121 YSLGADDLTPFGEQELVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFTIEGFOST 180

QY 181 KLADPGANPHQASPVNIIIEGAGYNTLDHGLCTAFESSELDGDDVEANFTAVFAPPIR 240
DB 181 KLKDPRAQPGSSPKIDVVISSEASSNNTLDPGCTVFESSELDGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPFDVTARTSDATLSPDCDLTFDHEWIOXDYDLSLG 300
DB 241 QLENDLSGVLTDEVTYLMDCMSFDTISTSTVDTKLSFPCDLTFDHEWIDYDLSLG 300

QY 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 KYGAGNPLGPAQGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNTWVSIFFALGNGTKPLSTTSVESIEETDGYSAWVTFPAAYVEMMQEAEKEP 420
DB 361 HDNGIISILFALGNGTKPLSTTTVENITOTDGFSSAWVTFPAAYVEMMQEAEKEP 420

QY 421 LVRVLNDRVPLHCGVDKLGCRKRDDEVEGLSFARSGGNWCECA 467
DB 421 LVRVLNDRVPLHCGVDKLGCRKRDDEVEGLSFARSGGNWCECA 467

```

RESULT 2

```

PHYA_ASPNG STANDARD; PRT; 467 AA.
AC P34752;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).
GN PHYA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93252284; PubMed=8387447;
RA van Hartingsveldt W., van Zeijl C.M.J., Harteveld G.M.,
RA Gouka R.J., Suykerbuyk M.E.G., Luiten R.G.M., van Paridon P.A.,
RA Selden G.C.M., Veenstra A.E., van Gorcom R.F.M.,
RA van den Hondel C.A.M.J.J.;

```

```

RT "Cloning, characterization and overexpression of the phytase-encoding
RT gene (phyA) of Aspergillus niger.";
RN Gene 127:87-94(1993).
RP SEQUENCE FROM N.A.
RA Mullaney E.J.;
RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 24-464.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=93249451; PubMed=8387289;
RA Ullah A.H.J., Dischinger H.C. Jr.;
RT "Aspergillus ficum phytase: complete primary structure elucidation
RT by chemical sequencing.";
RL Biochem. Biophys. Res. Commun. 192:747-753(1993).
RN [4]
RP SEQUENCE OF 71-93.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=91298982; PubMed=1648914;
RA Ullah A.H.J., Cummins B.J., Dischinger H.C. Jr.;
RT "Cyclohexanedione modification of arginine at the active site of
RT Aspergillus ficum phytase.";
RL Biochem. Biophys. Res. Commun. 178:45-53(1991).
RN [5]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=89160685; PubMed=2852807;
RA Ullah A.H.J.;
RT "Aspergillus ficum phytase: partial primary structure, substrate
RT selectivity, and kinetic characterization.";
RL Prep. Biochem. 18:459-471(1988).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC STRAIN=NRRL 3135 / VAN TIEGHEM / FICUUM;
RX MEDLINE=97307250; PubMed=9164457;
RA Kostrewa D., Gruning-Leitch F., D'Arcy A., Broger C., Mitchell D.,
RA van Loon A.P.;
RT "Crystal structure of phytase from Aspergillus ficum at 2.5-A
RT resolution.";
RL Nat. Struct. Biol. 4:185-190(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -!- BIOTECHNOLOGY: IS USED AS A FOOD AND FEED ADDITIVE. IT CAN
CC FACILITATE THE DEGRADATION OF PHYTIN IN SOYBEAN AND OTHER SEEDS
CC USED AS FOOD FOR MONOGASTRIC ANIMALS. SOLD BY NOVO NORDISK UNDER
CC THE NAME PHYTASE NOVO.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z16414; CAA78904.1; -
DR EMBL; M94550; AAA32705.1; -
DR PIR; JN0482; JN0482.
DR PIR; JN0023; JN0023.
DR PIR; JN0656; JN0656.
DR PDB; 1IHP; 18-MAR-98.
DR InterPro; IPR000560; -.
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolase; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 467 3-PHYTASE A.
FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).

```

```

FT DISULFID 31 40
FT DISULFID 71 414
FT DISULFID 215 465
FT DISULFID 264 282
FT DISULFID 436 444
FT CARBOHYD 27 27
FT CARBOHYD 59 59
FT CARBOHYD 105 105
FT CARBOHYD 120 120
FT CARBOHYD 207 207
FT CARBOHYD 230 230
FT CARBOHYD 339 339
FT CARBOHYD 352 352
FT CARBOHYD 376 376
FT CARBOHYD 388 388
SQ SEQUENCE 467 AA; 51086 MW; 88FE8F3584341D6D CRC64;

Query Match 74.9%; Score 1848; DB 1; Length 467;
Best Local Similarity 74.3%; Pred. No. 4.5e-140;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSLGLAVPASRNQSSCDTVDGYQCFSETSHLWGYAPFFSLANE 60
Db 1 MGVSALLPLLYLLSGVTSLGLAVPASRNQSSCDTVDGYQCFSETSHLWGYAPFFSLANE 60

QY 61 SATSPDVPKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SATSPDVPKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQOMVNSGIFKRYRYKALARKIVPFRASGSDRVIASAEKFTGFSQA 180
Db 121 YTLGADDLTPFGQOMVNSGIFKRYRYKALARKIVPFRASGSDRVIASAEKFTGFSQA 180

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLDEVDVNLMDMCPDPTVARTSDATOLSPFCDLFTDHEWIDYDYLQSLG 300
Db 241 ARLEAHLPGVNLDEVDVNLMDMCPDPTVARTSDATOLSPFCDLFTDHEWIDYDYLQSLG 300

QY 301 KYGAGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGAGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNMTWISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMOCEAKEP 420
Db 361 HDNMTWISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMOCEAKEP 420

QY 421 LRVVLNDRVPLRGCGVDKLGCRKDDFVEGLSFARSGCNWEECA 467
Db 421 LRVVLNDRVPLRGCGVDKLGCRKDDFVEGLSFARSGCNWEECA 467

```

```

RESULT 3
ID PHYB_EMENI
AC 000033;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 3-PHYTASE B PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE
DE 3-PHOSPHODIOLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHODIOLASE B).
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98007872; PubMed=9349716;
RA Pasamontes L., Haiker M., Henriquez-Huecas M., Mitchell D.B.,

```

```

RA van Loon A.P.;
RT "Cloning of the phytases from Emericella nidulans and the
RL thermophilic fungus Talaromyces thermophilus.";
RL Biochim. Biophys. Acta 1353:217-223(1997).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
CC FROM PHYTATE.
CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O =
CC D-MYO-INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + ORTHOPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U59803; AAB96871.1; -
CC InterPro: IPR000560; -
CC Pfam: PF00328; acid_phosphat_1;
CC PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 463
FT ACT_SITE 80 80
FT ACT_SITE 81 81
FT ACT_SITE 81 81
FT ACT_SITE 357 357
FT CARBOHYD 26 26
FT CARBOHYD 41 41
FT CARBOHYD 103 103
FT CARBOHYD 118 118
FT CARBOHYD 203 203
FT CARBOHYD 226 226
FT CARBOHYD 331 331
FT CARBOHYD 335 335
FT CARBOHYD 372 372
SQ SEQUENCE 463 AA; 51786 MW; ECC5827D1E1C82A2 CRC64;

Query Match 73.1%; Score 1804.5; DB 1; Length 463;
Best Local Similarity 72.5%; Pred. No. 1.3e-136;
Matches 338; Conservative 48; Mismatches 75; Indels 5; Gaps 3;

QY 1 MGVSALLPLLYLLSGVTSLGLAVPASRNQSSCDTVDGYQCFSETSHLWGYAPFFSLANE 60
Db 1 MAFETVALSLYLLSRVSAQA--PVVQNHSCNTADGGYQCFPNVSHVWGQYSPYFSEIQE 58

QY 61 SATSPDVPKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 59 SAISEDVPHGCEVTFVQVLSRHGARYPTSSASKAYSALIEAIQKNATSFVWGQYAFLESYN 118

QY 121 YTLGADDLTPFGQOMVNSGIFKRYRYKALARKIVPFRASGSDRVIASAEKFTGFSQA 180
Db 119 YTLGADDLTPFGQOMVNSGIFKRYRYKALARKIVPFRASGSDRVIASAEKFTGFSQA 178

QY 181 KLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 179 QLHDHGSK--RATPVVNVIIPEIDGPNNTLDHSCVSENDERADETEANFTAIMGPPIR 236

QY 241 ARLEAHLPGVNLDEVDVNLMDMCPDPTVARTSDATOLSPFCDLFTDHEWIDYDYLQSLG 300
Db 237 KRLENDLPGLIKLTNENVIYLMDCSFDTMARTAGTSLSPFCALFTTEKEWLDYDLSLS 296

QY 301 KYGAGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 297 KYGAGAGNPLGPAQGVGFVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 356

QY 361 HDNMTWISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMOCEAKEP 420
Db 357 HDNMTWISIFFALGLYNGTKPLSTTSVESIEETDGYASWTVFPAARAYVEMMOCEAKEP 415

```


the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U00029; AAB69729.1; -
 PIR; S48996; S48996.
 HSP; P34752; 1HP.
 SGD; S0001258; PHO12.
 InterPro; IPR000560; -
 Pfam; PF00328; acid_phosphat; 1.
 PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 Hydrolase; Glycoprotein; Signal; Multigene family.
 SIGNAL 1 17
 BY SIMILARITY.
 FT CHAIN 18 467
 ACT_SITE 75 75
 ACT_SITE 337 337
 CARBOHYD 97 97
 CARBOHYD 162 162
 CARBOHYD 192 192
 CARBOHYD 250 250
 CARBOHYD 315 315
 CARBOHYD 356 356
 CARBOHYD 390 390
 CARBOHYD 439 439
 CARBOHYD 445 445
 CARBOHYD 461 461
 SEQUENCE 467 AA; 52699 MW; BEC606CDF39B845B CRC64;

Query Match 16.0%; Score 394; DB 1; Length 467;

Best Local Similarity 25.7%; Pred. No. 5.4e-24; Mismatches 65; Indels 76; Gaps 14;

24 PRGNHSCDTPVGGYQCFFPEISHLWGTSPFFSLADESAISPDVPGKGRVTFVQVLSRHG 83
 21 PLGKLSIDIKICTEIPFL---GGSGPYYSFFGDYGISRLDPESECMQVQVGRHG 76
 84 ARPYTSSAKAYSAIEIAQKNAFKYAPLK-----TNYTL-----GADDL 128
 77 ERYPTVSKAKSINTTWYKLSNYTGFGSALFLNDYEFFIRDTKNLEMETTLANSVNL 136
 129 TPF-GQQMVNSGIFRYKALARKIVPF-TRASSDRVITASAEKFTGFSQAKLADPG 186
 137 NPYTGEMNAKRHARDELAQGYGVNENQTSFAVFTSNRCHDTAQYFDG-----LGDK- 190
 187 ANPHQASPVINVII-----PEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIRA 241
 191 -----FNISLQTISEAESAGANTLSAHSFCPAWDD-DVNDLKKYDKTGLSGIAK 240
 242 RLEAHLPGVNLDEDEVNLMDCPDTVARTSDATQLSPFCDLFTHDWQVYDIQSLGK 301
 241 RLKNEKGLNLSSANTFFAWCAYEINAR-----GYSDICNIETFKDELVRFSYGDQDET 295
 302 YYGAGNPLGPAQGVGFVNELIARLTSPVQDHTSTNHILDSNPATPLNATLYADFSH 361
 296 YQQTGPGYDVRSVCANLENASVKLLKSEVQDQ-----KWLSETH 337
 362 DNTWVSIFALLGNGTKPLSTTSVESIEETDGYSGASWTVPFAARAYVEMMOCEAKEPL 421
 338 DTDILNLTFTIGIIDQNNLTAEHPFMENT--FHRSWVPOGARVYTEKFCQ--SNDTY 393
 422 VRVLNDRVPLHGCVKDLGRCRDDR-----VEGLSFAR 457
 394 VRVINDAVVPIETCTGPGFSCINDFGYAEKRVAGTDFLK 436

RESULT 7

PPA3_YEAST

ID PPA3_YEAST

AC P24031;

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CONSTITUTIVE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO3 OR YBR092C OR YBR0813.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=85037940; PubMed=6093051;
 RA Bajwa W., Meyhack B., Rudolph H., Schweingruber A.-M., Hinnen A.;
 RT "Structural analysis of the two tandemly repeated acid phosphatase
 genes in yeast."
 RL Nucleic Acids Res. 12:7721-7739(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95208357; PubMed=7900426;
 RA Mannhaupt G., Stucka R., Ehle S., Vetter I., Feldmann H.;
 RT Analysis of a 70 kb region on the right arm of yeast chromosome II."
 RL Yeast 10:1363-1381(1994).
 CC -1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

EMBL; X01080; CAA35557.1; -
 EMBL; X78993; CAA55597.1; -
 EMBL; Z35961; CAA85045.1; -
 PIR; S44674; PABYCC.
 HSP; P34752; 1HP.
 SGD; S0000296; PHO3.
 InterPro; IPR000560; -
 Pfam; PF00328; acid_phosphat; 1.
 PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 Hydrolase; Glycoprotein; Signal; Multigene family.
 SIGNAL 1 17
 POTENTIAL.
 FT CHAIN 18 467
 ACT_SITE 75 75
 ACT_SITE 337 337
 CARBOHYD 97 97
 CARBOHYD 103 103
 CARBOHYD 162 162
 CARBOHYD 192 192
 CARBOHYD 250 250
 CARBOHYD 315 315
 CARBOHYD 356 356
 CARBOHYD 390 390
 CARBOHYD 439 439
 CARBOHYD 445 445
 CARBOHYD 456 456
 CARBOHYD 461 461
 CONFLICT 219 221
 DED -> MKT (IN REF. 1).
 SEQUENCE 467 AA; 52776 MW; 05FBB80DEB41BOFF CRC64;

Query Match 15.8%; Score 391; DB 1; Length 467;

Best Local Similarity 27.0%; Pred. No. 9.4e-24;

Mismatches 117; Conservative 61; Mismatches 192; Indels 64; Gaps 13;

QY 49 GTYSPFSLADESAISPDVPGKGRVTFVQVLSRHGARYPTSSASKAYSALIEIAQKNATA 108

Db 42 GGAGPYFFPGDYGISRLDPESECMQVQVGRHGARYPTSSASKAYSALIEIAQKNATA 101

[illegible]

RESULT	8	PPAD_YEAST	STANDARD;	PRT;	468 AA.
ID	PPAD_YEAST				
AC	F52290;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	PROBABLE ACID PHOSPHATASE YD1024C PRECURSOR (EC 3.1.3.2).				
GN	YD1024C OR D2815.				
OS	Saccharomycetes cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
ON	NCBI_TaxID=4932;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=S288C;				
RA	Andre B., Visser S.; Urrestarazu L.;				
RL	Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.				
CC	-1- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O -> AN				
CC	ALCOHOL + ORTHOPHOSPHATE.				
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.				

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z48432; CAA89335.1; -;
EMBL; Z74072; CAA98583.1; -;
SGD; S0002182; YDL024C.
InterPro; IPR000560; -;
Pfam; PF00328; acid_phosphat; 1.
PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
Hydrolase; Glycoprotein; Signal; Multigene family.
SIGNAL 1 20 BY SIMILARITY.
CHAIN 21 468 PROBABLE ACID PHOSPHATASE YDL024C.
ACT_SITE 76 76 NUCLIOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 338 338 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 98 98 N-LINKED (GLCNAC...) (POTENTIAL).
FT

[illegible]

RESULT	9
PHYB_ASPNG	
ID	PHYB_ASPNG
AC	STANDARD;
DT	PRT; 479 AA.
DT	P34754;
DT	01-FEB-1994 (Rel. 28, Created)
DT	01-FEB-1994 (Rel. 28, Last sequence update)
DT	01-JUN-1994 (Rel. 29, Last annotation update)
DE	3-PHYTASE B PRECURSOR (EC 3.1.1.38) (MYO-INOSITOL-HEXAPHOSPHATE
DE	3-PHOSPHOHYDROLASE B) (3 PHYTASE B) (MYO-INOSITOL HEXAKISPHOSPHATE
DE	3-PHOSPHOHYDROLASE B).
GN	PHYB.
OS	Aspergillus niger.
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC	Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus
OC	NCBI_TaxID=5061;
RN	[1]
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 20-101; 133-141 AND 376-399.
RX	MEDLINE=93371452; PubMed=7916610;
RA	Ehrlich K.C., Montalbano B.G., Mullaney E.J., Dischinger H.C. Jr.,
RA	Ullrich A.H.J.;
RT	"Identification and cloning of a second phytase gene (phyB) from
RT	Aspergillus niger (flicum).";
RL	Biochem. Biophys. Res. Commun. 195:53-57(1993).

QY 108 AFKGYAFLKTYNYTL-----GADDLT-PF-GEQOMVNSGIKRYRYKAL--ARKIYVF 157
 Db 109 EYKGLAFLNDWYYPNENYNAETSPYAGLLDAYNHGNDYKARYGHLNGETVVPF 168
 QY 158 IRASGSDRVIAAEKIEGFSQAKLADPGANPHQASPVINVIIEGAGYNNTLDHGLCTA 217
 Db 169 F-SSGGRVIETARKEGEF-----FGVNSTNAALNII 201
 QY 218 FESEELG-----DDVEANETAVFAPPIR--ARLEAHLPGVNLDTDEDVNNLMD 262
 Db 202 SESEVMGADSLPTCDTNDQTTCDNLTVQLPQFVAAARLNSQPNGLNTASDVYNNLMV 261
 QY 263 MCFDFIVARTSDATQISPCDLFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFVNE 322
 Db 262 MASFELNAR-----PFSNMINAFTQDEWVSFGYVEDLNYIYCAGFGDKNMAVAGVYANA 316
 QY 323 LIARLTHSPVQDHTSNHTLSDNPATFPLNATLYADFSDHNTMVSIFTFALGLY--NGTKP 380
 Db 317 STLTLNQGPKE-----AGSLFFNFADHTNITPILAAALGVLIPIEDLP 358
 QY 381 LSTTSVESIEETDGYASWTVFAARAYVEMMOCEA-----EKEPLRVLVNDRVVPVPLHCC 436
 Db 359 L-----DRVAFGNPISIGNIVPMGCHLTIERLSQATLSDEGTYRVRLVNEAVLPFND 413

RESULT 11

PPAL_PICPA STANDARD; PRT; 468 AA.
 AC P52291;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE ACID PHOSPHATASE PHO1 PRECURSOR (EC 3.1.3.2).
 GN PHO1.
 OS Pichia pastoris (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4922;
 RN P1
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96001238; PubMed=7557473;
 RA Payne W.E., Gannon P.M., Kaiser C.A.;
 RT "An inducible acid phosphatase from the yeast Pichia pastoris:
 RT characterization of the gene and its product.";
 RL Gene 163:19-26(1995).
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE
 CC -!- INDUCTION: BY PHOSPHATE STARVATION.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U28658; AAA85503.1;
 DR HSSP; P34752; 11HP
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Hydrolase; Glycoprotein; Signal; Multigene family.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 468 ACID PHOSPHATASE PHO1.
 FT ACT_SITE 84 84 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 345 345 PROTON DONOR (BY SIMILARITY).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 468 AA; 52690 MW; AE555B8E27718C2C CRC64;

Query Match 14.2%; Score 350; DB 1; Length 468;
 Best Local Similarity 25.2%; Pred. No. 1.8e-20;

Matches 124; Conservative 79; Mismatches 20; Indels 90; Gaps 21;

QY 6 VLLSIATL--FGSTGTALGPRGNSHSCDVGGYQCFPEISHLMQTSYSPF--FSLADE 60
 Db 10 IILALATLQSVFAVELQHLVGLVNDPRYPQRT--DQYNILRLHGLG-GPYLYGNGWIAAE 67
 QY 61 SAISPDVPKCRVTFVOVLSRHGARYPTSSAKAYSALIE-AIOKNATAFKGYAFLKTY 119
 Db 68 SEI-----ESCTIDQAHLLMRHGERYPSTNVGKLEALYQKLLDADVEVPTGPLSEFFQDY 122
 QY 120 NYTLG---ADDLTPFGEQOMVNSGIK---YRYKAL-----ARKIVPFRASGSDR 165
 Db 123 DYFVSDAANTYEQETTKGYSGLNTAFDFTGLRERYDHLINTSEEGKL--SYWAGSQER 180
 QY 166 VIASAEKFIKGFOSAKLADPGANPHOASPVINVIPE---GAGYNNNTLDHGLCT-----A 217
 Db 181 VVDFAKYFAOGFMKSNYTD-----NVEVVAAEEKSGQLNSLTARISCPNYSNHI 230
 QY 218 FESEELGDDV---EANTAVFAPIRAREHLPGVNLDTDEDVNNLMDMCPDTVARTSD 274
 Db 231 YKDGDFPDNIAEREA-----RLNTLSPGFNITADDIPTIALYCGPELNVGRGE- 278
 QY 275 ATQLSPFCDLFTHDEWIQYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLTHSPVQD 334
 Db 279 -----SSFCDVLSREALITYAYLRDLGLVNVNGNPLGKTIGYVYAN----- 321
 QY 335 HTSTNHTLSDNPATFPLNATLYADFSDHNTMVSIFTFALGLYNTK-PLSTTSVESIEETD 393
 Db 322 --ATROLLENTEAD-PRDYPLYSFSDHDTLLQVFTSLGLFNVTDLPL-----DQIQFT 373
 QY 394 GYSASWTVFAARAYVEMMOCC--EAEKEPLRVLVNDRVVPVPLHGGCGVKLGRCKRDFVE 451
 Db 374 SFKSTEIVPMGARLLTERLLCTVEGEEKYVVRTILNDVAFPLSDCSGPGFSCPLNDYVS 433
 QY 452 GLSFARSGGNWEE 464
 Db 434 RLEALNEDSDFAE 446

RESULT 12

PPA5_KLULA STANDARD; PRT; 469 AA.
 AC P52289;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE REPRESSIBLE ACID PHOSPHATASE PRECURSOR (EC 3.1.3.2).
 GN PHO5.
 OS Kluyveromyces lactis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=28985;
 RN P1
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 2359/152;
 RA Perminan E.;
 RL Thesis (1995), University of Salamanca, Spain.
 CC -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
 CC ALCOHOL + ORTHOPHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- INDUCTION: REPRESSED BY INORGANIC PHOSPHATE (BY SIMILARITY).
 CC -!- PTM: GLYCOSYLATED DURING SECRETION ACROSS THE MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

1

CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; Z47072; CAA87370.1; -
DR EMBL; Z54342; CAA87370.1; JOINED.
DR EMBL; Z54342; CAA91156.1; -
DR EMBL; Z47072; CAA91156.1; JOINED.
DR WormPep; P26C11.1; CE05732.
DR InterPro; IPR000560; -
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hypothetical protein: Hydrolase.
FT ACT_SITE 356 356 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT ACT_SITE 399 399 BY SIMILARITY.
FT ACT_SITE 636 636 PROTON DONOR (BY SIMILARITY).
FT DISULFID 702 708 BY SIMILARITY.
SQ SEQUENCE 755 AA; 87783 MW; E7F3050115A235EB CRC64;

Query Match 5.8%; Score 142; DB 1; Length 755;
Best Local Similarity 23.1%; Pred. No. 0.0014;
Matches 94; Conservative 53; Mismatches 132; Indels 128; Gaps 22;
QY 57 LADESAISPDVFKGKRVFQVLSRHGRYPTSSASKAYSALIEAIOKNATAFKGYAFL 116
DB 332 LAIEKSFMEV-DNMKLEFVQTIWHDG-----SALEGLFPISEKNWTEGG----- 378
QY 117 KTYNTLGGADLTPEGOQVNSGKFYRRY-----KALARKIYVFTIRAGSDRVI 167
DB 379 -----GLGELTPMGSMNNLGTIFRRYVEDQOFLSHRYAAKEI--YIRSTNLNRTI 429
QY 168 ASAEKIEGFQSAKLADPGANPHQASPINVIPEGAGYNNLTLD-----HGL- 214
DB 430 ISAMSLLYG-----MFPFGA-----WNI---QGVDPNDVDVWQOQGTTFIPVHVDGID 473
QY 215 -----CTAFEE-----SELGDDVEANFTAVFAPPTRARLEAHLPGVNLTDDEVVNL 260
DB 474 QCAVAQLCMCRFEQLQEKWAEI-DEVKNATVAMIA--LNRVAAF-----YVWTDQ----- 523
QY 261 MDMCPEDTVARTSDA--TQLSPFCD-LFTHDEWIQYDYLQSLGKYG-----YGAGNP 310
DB 524 -----EKENRYTDAMKQCRNWFNDTMYQQLPWYNEDLYNEAQRTYAPFKRFTGNGNP 577
QY 311 -----LGAQGVGVNELIARLTHSPVDHTSTNHTLD-SNPATFPLNATLY 356
DB 578 KPSIVDGDIDPOEVSTLOGGPLLNEIFEREKIRCVADAENCSDIYKPLKF-----Y 631
QY 357 ADPSHDNTMVSIFFALGLYNGTKPLSTSVESIEETDGYASASWTVPF 403
DB 632 AYSSHDQVYALLVTLGI-----TDVVKVDGWDPTSSSLTIEY 670

Search completed: October 26, 2001, 15:18:05
Job time: 451 sec

OM protein - protein search, using sw model

Run on: October 26, 2001, 15:17:26 ; Search time 75.85 Seconds
(without alignments)
814.587 Million cell updates/sec

Title: US-09-488-265-31
Perfect score: 2468
Sequence: 1 MGFFVLLSLATLFGSTSGT.....DFVEGLSFARSGGNWEECPA 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1888	76.5	465	3 000092	000092 aspergillus
2	1847	74.8	467	3 090027	090027 aspergillus
3	1841	74.6	467	3 093838	093838 aspergillus
4	1828	74.1	467	3 093838	093838 aspergillus
5	1822	73.8	466	3 000100	000100 aspergillus
6	1752	71.0	466	3 000085	000085 aspergillus
7	1710	69.3	466	3 000096	000096 talaromyces
8	1334.5	54.1	487	3 000107	000107 thielavia h
9	338.5	13.7	469	3 093846	093846 kluyveromyc
10	334.5	13.6	442	3 074677	074677 pichia angu
11	311.5	12.6	463	3 060172	060172 schizosacch
12	204.5	8.3	460	5 093846	093846 kluyveromyc
13	196	7.9	451	11 035217	035217 rattus norv
14	196	7.9	467	5 096421	096421 drosophila
15	188	7.6	481	11 092216	092216 mus musculu
16	188	7.6	487	4 095172	095172 homo sapien
17	187	7.6	487	4 09UNW1	09UNW1 homo sapien
18	185	7.5	487	4 09UGA3	09UGA3 homo sapien
19	181.5	7.4	453	5 096420	096420 drosophila

20	180.5	7.3	453	5 09W438	09W438 drosophila
21	177.5	7.2	449	13 092170	092170 gallus gall
22	168.5	6.8	198	3 09UTX1	09UTX1 schizosacch
23	149.5	6.1	468	10 004509	004509 arabidopsis
24	144	5.8	274	11 093JD5	093JD5 mus musculu
25	130	5.3	449	5 019076	019076 caenorhabdi
26	128.5	5.2	374	11 09JMG5	09JMG5 mus musculu
27	127.5	5.2	381	11 09QXG5	09QXG5 mus musculu
28	124.5	5.0	513	2 046334	046334 comamonas t
29	118	4.8	380	5 022525	022525 caenorhabdi
30	113.5	4.6	381	11 09QXH7	09QXH7 mus musculu
31	112.5	4.6	513	2 024719	024719 comamonas t
32	110	4.5	683	5 000838	000838 leishmania
33	110	4.5	707	5 000839	000839 leishmania
34	109.5	4.4	730	5 020826	020826 caenorhabdi
35	108.5	4.4	758	4 09P2C1	09P2C1 homo sapien
36	108.5	4.4	827	2 047871	047871 eubacterium
37	107	4.3	354	5 019709	019709 caenorhabdi
38	106	4.3	763	2 09R6Y7	09R6Y7 anabaena sp
39	103.5	4.2	1013	14 09ILX9	09ILX9 retroperito
40	103	4.2	421	4 09UIG6	09UIG6 homo sapien
41	103	4.2	428	4 09NPH0	09NPH0 homo sapien
42	103	4.2	452	5 019175	019175 caenorhabdi
43	102.5	4.2	1225	5 020330	020330 caenorhabdi
44	101.5	4.1	642	1 028446	028446 methanobact
45	101	4.1	408	5 019461	019461 caenorhabdi

ALIGNMENTS

RESULT 1

000092 ID 000092 PRELIMINARY: PRT: 465 AA.

AC 000092;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-
DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
DE PHOSPHOHYDROLASE A).

GN PHYA.

OS Aspergillus fumigatus (Sartorya fumigata),
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_Taxid=5085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-37.
RC STRAIN-ATCC 34625;
RX MEDLINE-97288063; PubMed-9143104;
RA Pasamontes L., Haiker M., Wyss M., Tessier M., van Loon A.P.G.M.;
RT "Gene cloning, purification, and characterization of a heat-stable
Phytase from the fungus Aspergillus fumigatus";
RL Appl. Environ. Microbiol. 63:1696-1700(1997).
CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
FROM PHYTATE.
CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
INGOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- MISCELLANEOUS: SHOWS HIGH ACTIVITY WITH 4-NITROPHENYL PHOSPHATE AT
A PH RANGE OF 3 TO 5 AND WITH PHYTIC ACID AT A PH RANGE OF 2.5 TO
7.5. IT IS ABLE TO WITHSTAND TEMPERATURES UP TO 100 DEGREES
CC CELSIUS OVER A PERIOD OF 20 MIN, WITH A LOSS OF ONLY 10% OF THE
CC INITIAL ENZYMATIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL; U59804; AAB96872.1; -;
DR HSSP; P34752; 1IHP.
DR InterPro; IPR000560; -;
DR Pfam; PF00328; acid_phosphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase; Glycoprotein; Signal.
FT SIGNAL 1 26

FT CHAIN 27 465 3-PHYTASE A.
 FT ACT_SITE 81 81 REQUIRED FOR BINDING SUBSTRATE (BY
 FT SIMILARITY).
 FT ACT_SITE 82 82 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 359 359 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 30 39 BY SIMILARITY.
 FT DISULFID 70 412 BY SIMILARITY.
 FT DISULFID 213 463 BY SIMILARITY.
 FT DISULFID 262 280 BY SIMILARITY.
 FT DISULFID 434 442 BY SIMILARITY.
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 465 AA; 50836 MW; 86FC1D9058C9B2C9 CRC64;

Query Match 76.5%; Score 1888; DB 3; Length 465;
 Best Local Similarity 76.9%; Pred. No. 4.5e-144;
 Matches 359; Conservative 35; Mismatches 71; Indels 2; Gaps 2;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNSHSCDVTGQYQCFPEISHLMGTSPFFSLADE 60
 DB 1 MYTLFLLSAAYLLSGRVSAAPSSAG-SKSCDVTDLGQYQCFPEISHLMGTSPFFSLEDE 59
 QY 61 SAISPDPVKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFELKTYN 120
 DB 60 LSVSSKLPKDCRTTLVQVLSRHGARYPTSSSKSKYKLVTAIQANATDFKGFALKTYN 119
 QY 121 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 120 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 179
 QY 181 KLADPGANPHQASPVINVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPTR 240
 DB 180 KLADPGAT-NRAAPALSVIIPESITFNNTLDHGVTCKPEASQGLGDEVAANFTALFAPDIR 238
 QY 241 ARLEAHLPGVNLTDDEVDVNLMDMCPDVTARTSDATQSPFCDLTFHDEWIDYDYLQSLG 300
 DB 239 ARAEKPLPGVTLTDEVDVNLMDMCPDVTARTSDATQSPFCDLTFHDEWIDYDYLQSLG 298
 QY 301 KYCYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 299 KYCYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 358
 QY 361 HDNTMVSIFPAGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMOCEAKEP 420
 DB 359 HDNSMVSIFPAGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMOCKSEKP 418
 QY 421 LVRVLNDRVVPVPLHSGCVKGLGRCKRDDVFEGLSFARSGGNWEECEFA 467
 DB 419 LVRVLNDRVVPVPLHSGCVKGLGRCKRDDVFEGLSFARSGGNWEECEFS 465

RESULT 2
 Q90U27

ID Q90U27 PRELIMINARY; PRT; 467 AA.
 AC Q90U27;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE MYO-INOSITOL HEXAPHOSPHATE PHOSPHOHYDROLASE PRECURSOR (EC 3.1.3.8).
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RA Hongning W., Qi W., Jing X.;
 RP "PCR, cloning and characterization of the phytase (phyA) gene of

RT Aspergillus niger (China Strain).";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF218813; AAF25481.1; --
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; --
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 KW Signal; Lyase; Hydrolase.
 FT SIGNAL 1 19 POTENTIAL.
 SQ SEQUENCE 467 AA; 51029 MW; F4300A8F165EBF92 CRC64;

Query Match 74.8%; Score 1847; DB 3; Length 467;

Best Local Similarity 74.1%; Pred. No. 9.2e-141;
 Matches 346; Conservative 45; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNSHSCDVTGQYQCFPEISHLMGTSPFFSLADE 60
 DB 1 MGVSAYLLPLYLSGVTSGLAVPASRNQSTCDVTGQYQCFSETSHLWGOYAPFFSLANE 60
 QY 61 SAISPDPVKGRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFELKTYN 120
 DB 61 SAISPDPVACRCVTFQAQLSRHARVPTDSKGYKALIEEQONATTFDGKAYFLKTYN 120
 QY 121 YTLGADDLTPFGEQVQVNSGKIFRYRYKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 DB 121 YSLGADDLTPFGEQELVNSGKIFRYQYVESLTRNIIPFIRSSGSRVIAASKKPIEQFQST 180
 QY 181 KLADPGANPHQASPVINVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPTR 240
 DB 181 KLKDPAPQPCQSSPKIDVISESSNNLTDPGTCVTFEDELADAVANFTATFVPTIR 240
 QY 241 ARLEAHLPGVNLTDDEVDVNLMDMCPDVTARTSDATQSPFCDLTFHDEWIDYDYLQSLG 300
 DB 241 QLENDLSGSLTDEVTYLMDCSFDTISTSTVDIKLSPFCDLTFHDEWIDYDYLQSLK 300
 QY 301 KYCYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 KYCYGAGNPLGPAQGVGNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 QY 361 HDNTMVSIFPAGLYNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMOCEAKEP 420
 DB 361 HDNGIISILFALGYNKTKPLSTTIVQNTITQDGFSSAWTVPFASRLYVEMMOCEAKEP 420
 QY 421 LVRVLNDRVVPVPLHSGCVKGLGRCKRDDVFEGLSFARSGGNWEECEFA 467
 DB 421 LVRVLNDRVVPVPLHSGCPDALGRCTRDSFVKGLSFARSGGDWAECEFA 467

RESULT 3
 Q93838

ID Q93838 PRELIMINARY; PRT; 467 AA.
 AC Q93838;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE PHYTASE.
 GN PHVA.
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SK-57;
 RA Nagashima T., Kondo H., Anazawa H., Terasaki Y.;
 RT "Phytase having high-affinity for phytic acid."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB022700; CAB19824.1; --
 DR HSSP: P34752; 1IHP.
 DR InterPro: IPR000560; --
 DR Pfam: PF00328; acid_phosphat; 1.

DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
 SQ SEQUENCE 467 AA; 51028 MW; 7A38AD543EDC265C CRC64;

Query Match 74.1%; Score 1841; DB 3; Length 467;
 Best Local Similarity 73.2%; Pred. No. 2.8e-140;
 Matches 342; Conservative 49; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRNHSCTVDGGYQCFPEISHLWGTYPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNSQCTVDGGYQCFSETSHLWGYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSAKYASALIEAIOKNATPAFGKYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFAQVLSRHGARYPTDSKGKYSALIEEIQONATTPGKYAFKLTYN 120
 QY 121 YTLGADLLTPFGQOMVNSGKIFRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 YSLGADLLTPFGQELVNSGVKFYQYESLTRNIVPFFIRSSGSSRVIASGNKFIQFQST 180
 QY 181 KLADPCANPHQASPVNVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 DB 181 KLKDPRAQPGQSSPKIDVVISSEASTSNNTLDPGTCTVFEDSELADDEANFTATVP 240
 QY 241 ARLEAHLPGVNLTDVNVLMDCMCPDVTARTSDATQISLSPFCDLTFHDEWYQYDLSLG 300
 DB 241 ORLENDLSGVSLTDEVTYVLMDCMCFDITSTVDTKLSLSPFCDLTFHEWINYDLSLN 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYYGAGNPLGPTQCGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HNTWYISFFALGLYNGTKPLSTTSVESIEERTDGYASWTVPFAARAYVEMMOCEAKEP 420
 DB 361 HNGIISILFALGLYNGTKPLSTTAENITQTDGFSARTVPFAARMYVEMMQCQSEEP 420
 QY 421 LVRVLNDRVPLHCGVDKLGCRKRDDEFVGLSFARSGGNWEECF 467
 DB 421 LVRVLNDRVPLHCGVDALGRCTRDSEFVKGLSFARSGGDMGECFA 467

RESULT 4

ID Q9HEQ0 PRELIMINARY; PRT; 467 AA.
 AC Q9HEQ0;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PHYTASE.
 OS Aspergillus ficuum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang L., An L., Wang Y., Yuan X.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY013315; AA40885.1; -;
 SQ SEQUENCE 467 AA; 51012 MW; 3F69AD543C0B565B CRC64;

Query Match 74.1%; Score 1828; DB 3; Length 467;
 Best Local Similarity 73.0%; Pred. No. 3.2e-139;
 Matches 341; Conservative 49; Mismatches 77; Indels 0; Gaps 0;

QY 1 MGVEVLLSTATLFGSTSGTALGRNHSCTVDGGYQCFPEISHLWGTYPFFSLADE 60
 DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNSQCTVDGGYQCFSETSHLWGYAPFFSLANK 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSAKYASALIEAIOKNATPAFGKYAFKLTYN 120
 DB 61 SAISPDVPAGCHVTFAQVLSRHGARYPTDSKGKYSALIEEIQONATTPGKYAFKLTYN 120

QY 121 YTLGADLLTPFGQOMVNSGKIFRYKALARKIVPFFIRASGSDRVIASAEKFIQFQSA 180
 DB 121 YSLGADLLTPFGQELVNSGVKFYQYESLTRNIVPFFIRSSGSSRVIASGNKFIQFQST 180
 QY 181 KLADPCANPHQASPVNVIIPGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 DB 181 KLKDPRAQPGQSSPKIDVVISSEASTSNNTLDPGTCTVFEDSELADDEANFTATVP 240
 QY 241 ARLEAHLPGVNLTDVNVLMDCMCPDVTARTSDATQISLSPFCDLTFHDEWYQYDLSLG 300
 DB 241 ORLENDLSGVSLTDEVTYVLMDCMCFDITSTVDTKLSLSPFCDLTFHEWINYDLSLN 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVQDHTSNHTLDSNPATFPLNATLYADFS 360
 DB 301 KYYGAGNPLGPTQCGVYANELIARLTHSPVDDTSSNHTLDSNPATFPLNSTLYADFS 360
 QY 361 HNTWYISFFALGLYNGTKPLSTTSVESIEERTDGYASWTVPFAARAYVEMMOCEAKEP 420
 DB 361 HNGIISILFALGLYNGTKPLSTTAENITQTDGFSARTVPFAARMYVEMMQCQSEEP 420
 QY 421 LVRVLNDRVPLHCGVDKLGCRKRDDEFVGLSFARSGGNWEECF 467
 DB 421 LVRVLNDRVPLHCGVDALGRCTRDSEFVKGLSFARSGGDMGECFA 467

RESULT 5

ID 000100 PRELIMINARY; PRT; 466 AA.
 AC 000100;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE PRECURSOR (EC 3.1.3.6) (MYO-INOSITOL-HEXAKISPHOSPHATE 3-
 DE PHOSPHOHYDROLASE) (3 PHYTASE) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE).
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CBS 116.46;
 RA Pasamontes L., Haiker M., Henriquez Huecas M., Hug D., Mitchell D.B.,
 RA Broger C., van Loon A.P.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHITATE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL: U60412; AAB58465.1; -;
 DR HSSP: P34752; I1HP.
 DR InterPro: IPR000560; -;
 DR Pfam: PF00328; acid_phosphat; 1.
 DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 466 3-PHYTASE.
 FT ACT_SITE 82 82 REQUIRED FOR BINDING SUBSTRATE (BY
 FT ACT_SITE 83 83 SIMILARITY).
 FT ACT_SITE 361 361 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT DISULFID 31 40 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).

Sat Oct 27 15:25:56 2001

us-09-488-265-31.rspt

FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51055 MW; F2AECECIAF7C22C4 CRC64;

Query Match 73.8%; Score 1822; DB 3; Length 466;
 Best Local Similarity 72.3%; Pred. No. 9,6e-139;
 Matches 337; Conservative 46; Mismatches 88; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNHSDCTVDRGYQCFPELHKWGLYAPYFSLQDE 60
 DB 1 MGVEVLLSIATLFGSTGTALGPRGNHSDCTVDRGYQCFPELHKWGLYAPYFSLQDE 60

QY 61 SAISDPVPGKRVTFVQVLSRGARYPTSSASKAYSALEIAQKNAATKGYAFLKTYN 120
 DB 61 SPPLDVPDCHTFVQVLSRGARYPTSSASKAYSALEIAQKNAATKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 YSMGSENLNPFGRNQLDGLGAQFYRYDTRHINPFVRAADSRVHSAEKFEVGFQNA 180

QY 181 KLADPGANPHQASPVNVIPIEGAGYNNITLDHGLCTAPESELGDDVEANFTAVFAPPIR 240
 DB 181 RQDDHHPHOPSPRVDAIPEGSAYNNITLDHGLCTAPESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVVDVNLMDCPDVTARTSDATQLSPFCDLTHDEWIQDYLSLG 300
 DB 241 KRLEADLPVQLSADDVNLMDCPDVTARTSDATQLSPFCDLTHDEWIQDYLSLG 300

QY 301 KYGYGAGNPLGPAQGVGVFVNEILARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPAQGVGVFVNEILARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

QY 361 HDNTWVSIFPAGLGLYNGTKPLSTTSVESTETDGYSAKTSKAYSALEIAQKNAATKGYAFLKTYN 420
 DB 361 HDNTWVSIFPAGLGLYNGTKPLSTTSVESTETDGYSAKTSKAYSALEIAQKNAATKGYAFLKTYN 420

QY 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDAFVAGLSFAQAGGNWADCF 466
 DB 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDAFVAGLSFAQAGGNWADCF 466

RESULT 6
 ID 000085 PRELIMINARY; PRT; 466 AA.
 AC 000085;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 13, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE A).
 CN PHVA.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A1;
 RX MEDLINE=97177792; PubMed=9025298;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
 RT "The phytase subfamily of histidine acid phosphatases: isolation of
 RT genes for two novel phytases from the fungi *Aspergillus terreus* and
 RT *Myceliophthora thermophila*.";
 RL Microbiology 143:245-252(1997).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF
 CC 2.5 TO 7.5, WITH MAXIMAL ACTIVITY AT PH 5.5. ALSO ACCEPT 4-
 CC NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO
 CC MORE ACIDIC PH VALUES.
 CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
 DR EMBL; U59805; AAB52507.1; -;
 DR HSSP; P34752; 1IHP.
 DR InterPro; IPR000560; -;
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase; Glycoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 466 REQUIRED FOR BINDING SUBSTRATE (BY
 FT ACT_SITE 82 82 SIMILARITY).
 FT ACT_SITE 83 83 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
 FT ACT_SITE 361 361 PROTON DONOR (BY SIMILARITY).
 FT DISULFID 31 40 BY SIMILARITY.
 FT DISULFID 71 414 BY SIMILARITY.
 FT DISULFID 215 465 BY SIMILARITY.
 FT DISULFID 264 282 BY SIMILARITY.
 FT DISULFID 436 444 BY SIMILARITY.
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 466 AA; 51093 MW; 21CDB559C96AB66 CRC64;

Query Match 71.0%; Score 1752; DB 3; Length 466;
 Best Local Similarity 69.3%; Pred. No. 4,2e-133;
 Matches 323; Conservative 56; Mismatches 87; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNHSDCTVDRGYQCFPELHKWGLYAPYFSLQDE 60
 DB 1 MGVEVLLSIATLFGSTGTALGPRGNHSDCTVDRGYQCFPELHKWGLYAPYFSLQDE 60

QY 61 SAISDPVPGKRVTFVQVLSRGARYPTSSASKAYSALEIAQKNAATKGYAFLKTYN 120
 DB 61 SPPLDVPDCHTFVQVLSRGARYPTSSASKAYSALEIAQKNAATKGYAFLKTYN 120

QY 121 YTLGADDLTPFGQQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 YTLGADDLTPFGQQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180

QY 181 KLADPGANPHQASPVNVIPIEGAGYNNITLDHGLCTAPESELGDDVEANFTAVFAPPIR 240
 DB 181 RQDDHHPHOPSPRVDAIPEGSAYNNITLDHGLCTAPESELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDVVDVNLMDCPDVTARTSDATQLSPFCDLTHDEWIQDYLSLG 300
 DB 241 KRLEADLPVQLSADDVNLMDCPDVTARTSDATQLSPFCDLTHDEWIQDYLSLG 300

QY 301 KYGYGAGNPLGPAQGVGVFVNEILARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 KYGYGAGNPLGPAQGVGVFVNEILARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

QY 361 HDNTWVSIFPAGLGLYNGTKPLSTTSVESTETDGYSAKTSKAYSALEIAQKNAATKGYAFLKTYN 420
 DB 361 HDNTWVSIFPAGLGLYNGTKPLSTTSVESTETDGYSAKTSKAYSALEIAQKNAATKGYAFLKTYN 420

QY 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDAFVAGLSFAQAGGNWADCF 466
 DB 421 LVRVLVNDRVNPLHGCPTDKLGRCKRDAFVAGLSFAQAGGNWADCF 466

RESULT 7
 ID 000096 PRELIMINARY; PRT; 466 AA.
 AC 000096;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 13, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE 3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3-
 DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE
 DE PHOSPHOHYDROLASE A).
 CN PHVA.
 OS Aspergillus terreus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=33178;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A1;
 RX MEDLINE=97177792; PubMed=9025298;
 RA Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L., Loon A.P.;
 RT "The phytase subfamily of histidine acid phosphatases: isolation of
 RT genes for two novel phytases from the fungi *Aspergillus terreus* and
 RT *Myceliophthora thermophila*.";
 RL Microbiology 143:245-252(1997).
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE
 CC FROM PHYTATE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO-
 CC INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.

QY	239	IRARLEAHLPCVNLITDEDVNVLMDMCPFDIVARTSDATQLSPPCDLFTHDENIQYDYLQS	298
Db	236	IILEIKDHLPGVDLAVSDVPLMDLCPFFETIARNHTDT-LSPFCASTLQEQWQAVDYQS	294
QY	239	LGKYYGYGAGNPLCPAGCGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYAD	358
Db	295	LKYYGNGGNGPLCPAGCGVFNELIARMTHTSPVQDYITVNNHTLDSNPATFPLNATLYAD	354
QY	359	FSHDNTMYSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMOCEAEK	418
Db	355	FSHDNTMTSIFAALGLYNGTAKLSTTEIKSIEETDGYASAANTVPFGGRAYIEMMQCDDSD	414
QY	419	EPLVRLVNDRVVPLHCGGYDKLGRCKRDDFVEGLSFARSGGNWECFA	467
Db	415	EPVRLVNDRVVPLHGEVDLSGRCKRDDFVRGLSFARSGGNWEGCYA	463
RESULT	8		
ID	000107	PRELIMINARY;	PRT; 487 AA.
AC	000107;		
DT	01-JUL-1997	(TReMBLrel. 04, Created)	
DT	01-JUL-1997	(TReMBLrel. 04, Last sequence update)	
DT	01-MAY-2000	(TReMBLrel. 13, Last annotation update)	
DE	3-PHYTASE A PRECURSOR (EC 3.1.3.8) (MYO-INOSITOL-HEXAPHOSPHATE 3- DE PHOSPHOHYDROLASE A) (3 PHYTASE A) (MYO-INOSITOL HEXAKISPHOSPHATE GN. PHA.		
OS	Thielavia heterothallica.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
OC	Sordariales; Chaetomiaceae; Thielavia.		
OX	NCBI_TaxID=78579;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97177792; PubMed=9025298;		
RA	Mitchell D.B., Vogel K., Weimann B.J., Pasamontes L.,		
RA	van Loon A.P.G.M.;		
RT	"The phytase subfamily of histidine acid phosphatases: isolation of genes for two novel phytases from the fungi <i>Aspergillus terreus</i> and Myceliophthora thermophila.";		
RL	Microbiology 143:245-252(1997).		
CC	-1- FUNCTION: CATALYZES THE HYDROLYSIS OF INORGANIC ORTHOPHOSPHATE FROM PHYTATE.		
CC	-1- CATALYTIC ACTIVITY: MYO-INOSITOL HEXAKISPHOSPHATE + H(2)O = D-MYO- INOSITOL 1,2,4,5,6-PENTAKISPHOSPHATE + PHOSPHATE.		
CC	-1- SUBCELLULAR LOCATION: SECRETED.		
CC	-1- MISCELLANEOUS: SHOWS ACTIVITY WITH PHYTIC ACID AT A PH RANGE OF 3.5 TO 8.5, WITH MAXIMAL ACTIVITY BETWEEN PH 5.5 AND 6.0. ALSO ACCEPT 4-NITROPHENYL PHOSPHATE AS SUBSTRATE WITH A PH OPTIMA SHIFTED TO MORE ACIDIC PH VALUES.		
CC	-1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.		
DR	EMBL; U59806; AB52308.1; -		
DR	HSSP; P34752; 1IHP.		
DR	InterPro; IPR000560;		
DR	Pfam; PF00328; acid_phosphat; 1.		
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.		
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.		
KW	Hydrolase; Glycoprotein; Signal.		
FT	SIGNAL 1 ?		
FT	CHAIN ? 487		
FT	DOMAIN 267 270		
FT	DOMAIN 423 433		
FT	ACT_SITE 75		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID 26 35		
FT	DISULFID 64 421		
FT	DISULFID 208 485		
FT	DISULFID 260 289		
FT	DISULFID 456 464		
FT	CARBOHYD 165 165		
FT	CARBOHYD 200 200		
FT	ACT_SITE 76		
FT	ACT_SITE 368		
FT	DISULFID		

Sat Oct 27 15:25:56 2001

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulev G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong E., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AF003527; AAF49450.1;
DR FlyBase: FBgn0026061; Mippl.
DR InterPro: IPR000560;
DR Pfam: PF00328; acid_phosphat; 1.
SQ SEQUENCE 460 AA; 52685 MW; F00B25718E40807D CRC64;

Query Match 8.3%; Score 204.5; DB 5; Length 460;
Best Local Similarity 23.9%; Pred. No. 2.7e-08;
Matches 106; Conservative 64; Mismatches 171; Indels 103; Gaps 22;

QY 66 DVPK-----GCRVTFVQVLSRHGARYTSS-----ASKAYSALIEAIQKNATAFKSK----- 112
DB 46 DIDKQVLPGCCQPKQKWIHRHGTLPKSKMINKASRV-AELRDLIINNYQVARTKPTD 104
QY 113 -----YAFKYNVYTLGADDTLPGEQMVNSGKIFRYRYKALARKIVP-FIRASGSD- 164
DB 105 ALCOTDLIAIKLWKN-----SSITPDMEELYTAQGYEDLRGTAKLYQRYPTVLTANYNDT 161
QY 165 ---RVTSAEKLEGG-FQSAKLADPGANPHO-----ASPVINV-IIPGAGYNNLT 210
DB 162 YQRTTESKFAEAGLFGSQNAHVEIPKQDILLRPPDYCSFKNVNVKDEGSEYK-- 219
QY 211 DHGLCTAFEESELGDDVEANFTAVFAPPVIRARLEALPG--VNLTDEDVNNLMDMCPEDT 268
DB 220 -----FHOSKLYNTLAD-----ISTRL-----GFLVTLLEADIKLMDMCRVE- 258
QY 269 VARTSATQLSPCDLFTDHWIQTQYDYLQSLGKYGYGAGNPLGPAQGVGVNELLARLT 328
DB 259 --QAWNVDNRNSVMCGAFLEQITVFYLEDLYKYGGYGFPEPNAHLNRLVODLLTHLS 316
QY 329 HSPVQDHTSNHTLDSNPATFPLNATFLYADESHDNTMVSIFFAALGNYGKTKPLSTTSVES 388
DB 317 -NPVSPH-----VVAHFGHSTGLLLTLLTALGKQDKDIKLADNYDS 356
QY 389 IETDGYASWVYFAARAYVEMMOCEA--EKEPLRVLVNDRVPLHCGGVKDLGRCKR 446
DB 357 L-TSRRWKSLIDPFAANFVAVKYDCPADLDREKVV-FELNQAVQLDWCVS---GLCKW 411
QY 447 DDFVE-----GLSFARSGG 460
DB 412 SDVLEKYTKTIADACGEYCYRTGG 435

RESULT 13
O35217
ID O35217 PRELIMINARY; PRT; 451 AA.
AC O35217;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JUN-1999 (TREMBLrel. 11, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HEPATIC MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE (EC 3.1.3.-)
DE (MIPPL).
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 11-451 FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=LIVER;
RX MEDLINE=98028556; PubMed=9359836;
RA Craxton A., Caffrey J.J., Burkhardt W., Safranay S.T., Shears S.B.;
RT "Molecular cloning and expression of a rat hepatic multiple inositol
polyphosphate phosphatase";
RL Biochem. J. 328:75-81(1997).
RN [2]
RP CATALYTIC ACTIVITY.
RX MEDLINE=91358435; PubMed=1653239;
RA Nogimori K., Hughes P.J., Glennon M.C., Hodgson M.E., Putney J.W. Jr.,
Shears S.B.;
RT "Purification of an inositol (1,3,4,5)-tetraakisphosphate 3-phosphatase
activity from rat liver and the evaluation of its substrate
specificity";
RL J. Biol. Chem. 266:16499-16506(1991).
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF INOSITOL 1,3,4,5-
TETRAKISPHOSPHATE TO INOSITOL 1,4,5-TRIPHOSPHATE AND INOSITOL
1,3,4,5,6-PENTAKISPHOSPHATE TO INOSITOL 1,4,5,6-TETRAKISPHOSPHATE
AND THEN TO INOSITOL 1,4,5-TRISPHOSPHATE. DEPHOSPHORYLATES
INOSITOL HEXAKISPHOSPHATE WITHOUT SPECIFICITY TOWARDS A PARTICULAR
PHOSPHATE GROUP.
CC -1- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM LUMEN.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN KIDNEY AND LIVER.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
DR EMBL: AF012714; AAC53453.1; --
DR InterPro: IPR000560; --
DR Pfam: PF00328; acid_phosphat; 2.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN-1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; FALSE_NEG.
KW Hydroxylase; Glycoprotein; Phosphorylation; Multifunctional enzyme.
FT ACT_SITE 59 NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
FT MOD_RES 65 65 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 146 146 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 176 176 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 201 201 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 218 218 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 391 391 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 412 412 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 435 435 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 447 447 PHOSPHORYLATION (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC..) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC..) (POTENTIAL).
FT VARIANT 3 3 P -> H.
FT CONFLICT 327 327 S -> P (IN AAC53453).
FT SITE 448 451 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 451 AA; 51592 MW; EB1C05512A03020B.CRC64;

Query Match 7.9%; Score 196; DB 11; Length 451;
Best Local Similarity 21.3%; Pred. No. 1.3e-07;
Matches 91; Conservative 72; Mismatches 187; Indels 78; Gaps 17;

QY 52 SPFTS-----LADESA--ISPDVPGK-CRVTFVQVLSRHGARYPTSSASKAYS 96
DB 14 SPYGTGTRYEDVNPWLLGDPVAPRDELLAGTCTPVQLVALRHGTRYPTTKIRKLR 73
QY 97 ALIEAIQKNATAFKGYAFLKTYNTLTGADDTLPFGQQMVNSGKIFRYRYKALARKIVP 156
DB 74 QLQGLLOTRESVDGGSRVAAALDQWPLWDD---WMDGQLVEKGRQMDRLALALFP 130
QY 157 --FIR-----ASGSDRVIAAEKFIQFQSAKLADPGANPHQASPVINVIPEGAG 205

Db 131 DLFCRENYGLRLITSSKRCVDSAAFLQGLW--QHYPGLPP-----PDVSDMECDPPR 184
 QY 206 YNNTL-----DHGLCTAFEESELGDDVEANFTAVF-----APPPEARLE-----AHLPG 249
 Db 185 VNDKLMRFEDH--CEKL-----TEVERNATALYHVEAEKTEGEMOTVLKKAATLOVPV 237
 QY 250 VNLTDEVDVNLMDMCPDFTVARTSDATOLSPFCDFLTHDEWIQYDLSGLKGYGYGAGN 309
 Db 238 NNLNADLIQVAFETCSFDLAIQGVH---SPWCDVFDVDAKVLVLENDLKYWKRSYGY 293
 QY 310 PLGPAQGVGVNLIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIF 369
 Db 294 AINSRSSCNLFQDIFLHLDAVEQKORSQ-----PVSSVILQFHAETLLPL 342
 QY 370 FALGLYNGTKPLSTTSVESIEETDGYSSASWTVPFAARAYVENMQCEAKEP---LVRVL 425
 Db 343 SLMGFKDEPLTAYNFEQVIRE-FRSGHIVPYASNLFVLYHCEDAQTPOEKFOIOML 401
 QY 426 VNRVVPL 433
 Db 402 LNEKVLPL 409
 RESULT 14
 ID O96421 PRELIMINARY; PRT; 467 AA.
 AC O96421;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE 1.
 GN MIPPI OR C04123.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046913; AAD02436.1;
 DR FlyBase; FBgn0026061; Mippl.
 DR InterPro; IPR000560;
 DR Pfam; PF00328; acid_phosphat; 1.
 SQ SEQUENCE 467 AA; 53602 MW; 1A0571B9958B2632 CRC64;

Query Match 7.9%; Score 196; DB 5; Length 467;
 Best Local Similarity 22.8%; Pred. No. 1.3e-07;
 Matches 103; Conservative 66; Mismatches 172; Indels 110; Gaps 21;

QY 66 DYPK-----GCRVTFVOVLSRHGARYPTSS-----ASKAYSALIEALOKNATAFKGK---- 112
 Db 46 DIDKQVLPVGPQCPQKQWIFHRHGTLPLPKSMINKASRV-AELRDLIINQYVARTKPTD 104
 QY 113 -----YAPLKTYNTLGADLTTPGEOQMVNSGKIFYRRYKALARKIVPFI----- 158
 Db 105 ALCOYDLIAIKLWKN---SSITPDMEVLTAGQVEDLGTAKLQRYVPTLTANYNDT 161
 QY 159 -----RAGSDRVIAAEKFIQF-FOSAKIADPGANPHQ-----ASFVINV-IIPEG 203
 Db 162 YYQFRHTDTQRTTFKFAEGLFGSQNAHPVEIPKQDILLRPDYDCSSFRKNVYKDEG 221
 QY 204 AGYNTYLDHGLCTAFEESELGDDVEANFTAVFAPPPEARLEAHLPG--VNLTDEVDVNL 261
 Db 222 SEYK-----FHQSKLYNDTLAD-----ISTRN-----GFLYTLLEADIKLMY 259
 QY 262 DMCPTDVTARTSDATOLSPFCDFLTHDEWIQYDLSGLKGYGYGAGNPLGPAQGVGFN 321
 Db 260 DMCRYE---QAWNVRNVSWCAGAFLEQITVFEYLEDLKYGYGYPFENAHNCRVLQ 316

QY 322 ELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFPALGNGTKPL 381
 Db 317 DLTHLUS-NPVSPH-----VVXHFHSTGLLTLLTALGXKODIKL 356
 QY 382 STTSVESIEETDGYSSASWTVPFAAR--AYVENMQCEAKEPVLVRVLENDVWPLHGGVD 439
 Db 357 RADNYDSL-TSRWKSLSLIDPFAANFVAVKYDLPADLDREKVV-FELNOOAVOLDWCSV- 413
 QY 440 KIGRCRKRDFVE-----GLSFARSGG 460
 Db 414 --GLCKWSOVLEKYKTIADADCGEYICRTGG 442
 RESULT 15
 ID Q922L6 PRELIMINARY; PRT; 481 AA.
 AC Q922L6;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE MULTIPLE INOSITOL POLYPHOSPHATE PHOSPHATASE.
 GN MINPPI OR MIPP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129;
 RA Chi H., Wang J., Romano P.R., O'keefe R.J., Puzas J.E., Rosier R.N.,
 RA Reynolds P.R.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046908; AAD02434.1;
 DR MGD; MGI:1336159; Minppl.
 DR InterPro; IPR000560;
 DR InterPro; IPR000886;
 DR Pfam; PF00328; acid_phosphat; 2.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
 SQ SEQUENCE 481 AA; 54537 MW; 7CE78356B11F4606 CRC64;

Query Match 7.6%; Score 188; DB 11; Length 481;
 Best Local Similarity 20.8%; Pred. No. 6.2e-07;
 Matches 99; Conservative 70; Mismatches 196; Indels 112; Gaps 20;

QY 11 AILFGSTGTALPGRNHSCDVTGCGYOCFFPEISHLWGTYSPPFSLAD--ESA-----I 63
 Db 21 AALLSFARCSLPGRGD-----PVASVL---SPYFGTKTRYEDANPWLIV 62
 QY 64 SPDVPK-----GCRVTFVOVLSRHGARYPTSSASKAYSALIEALOKNATAFKGYAF 115
 Db 63 DPVAPRRDPELLAGTCTPVQVVALIRHGTRYPTTKIRKLQLOGGLQTRSDGGSQVA 122
 QY 116 LKTYNTLGADLTTPGEOQMVNSGKIFYRRYKALARKIVP--FIR-----ASGSD 164
 Db 123 AALAEWPLWYGD---WMDQGLVEKGRQDMRQLALRALAALFPDLFSRENVDRLRLITSSKH 179
 QY 165 RVIASAEKIEFGFSQAKLADPGANPHQAS-----PVINVIPEGAGYNTLDHGLCTAF 218
 Db 180 RCVDSSAFLQGLW--QHYHPLPPPDVSDMECGPPRINDKL-----MRFFDH--CEKF 229
 QY 219 EESELGDDVEANFTAVFAPPPEARLEAHLPG-----VNLTDEVDVNL- 260
 Db 230 L-----TDVERNETALY-----HVEAFKTPGEMQKVLKKAATLOVPVNSLNADIQVA 278
 QY 261 DMCPTDVTARTSDATOLSPFCDFLTHDEWIQYDLSGLKGYGYGAGNPLGPAQGVGF 320
 Db 279 FTFCSFDLAIKGVH---SPWCDVFDVDAKVLVLENDLKYWKRSYGYTINSRSSCNLF 334
 QY 321 NELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDHNTMVSIFPALGNGTKP 380
 Db 335 QDIFLHLDAVEQKORSQ-----PVSSPVLQFQGHAEITLPLLSLMLGYFKDKEP 383

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:04 ; Search time 57.24 Seconds
(without alignments)
494.609 Million cell updates/sec

Title: US-09-488-265-31

Perfect score: 2468

Sequence: 1 MGVEVLLSTATLFGSTSGT.....DFVEGLSFARSGGNWEECA 467

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2468	100.0	467	20	AA43170
2	2468	100.0	467	21	AA20527
3	2468	100.0	467	21	AA69569
4	2456	99.5	467	21	AA20534
5	2451	99.3	467	21	AA20533
6	2448	99.2	467	21	AA20524
7	2448	99.2	467	21	AA69566
8	2367	95.9	467	21	AA20532
9	2362	95.7	467	20	AA43169
10	2362	95.7	467	21	AA20526
11	2362	95.7	467	21	AA20531

12	2362	95.7	467	21	AA69568	Mutant phytase-1.
13	2322	94.1	441	21	AA20523	Consensus phytase
14	2314	93.8	467	20	AAW9382	Fungal consensus p
15	2308	93.5	467	20	AAW9380	Fungal phytase pro
16	2308	93.5	467	21	AAW9381	Fungal consensus p
17	2308	93.5	467	21	AA20515	Consensus phytase
18	2308	93.5	467	21	AA69558	Phytase-1, a conse
19	2307	93.5	467	21	AAW9383	Fungal consensus p
20	2305	93.4	467	20	AAW9384	Fungal consensus p
21	2299	93.2	467	20	AAW9385	Fungal consensus p
22	2298	93.1	467	20	AA69906	Ascomycete consens
23	2297	93.1	467	21	AA20530	Consensus phytase
24	2297	93.1	467	21	AA69572	Conspyl2, a deriv
25	2190	88.7	467	21	AA43172	Consensus phytase-
26	2190	88.7	467	21	AA69571	Phytase-7, a deriv
27	2187	88.6	467	21	AA20529	Consensus phytase
28	2182	88.4	441	21	AA20514	Consensus phytase
29	2093	84.8	437	21	AA20525	Consensus phytase
30	2093	84.8	437	21	AA69567	Phytase-11, a cons
31	2056	83.3	431	21	AA20535	Consensus phytase
32	2056	83.3	431	21	AA69557	Initial consensus
33	2046	82.9	467	21	AA69574	Mutant Aspergillus
34	2028.5	82.2	424	21	AA20536	Consensus phytase
35	2028.5	82.2	424	21	AA69565	Initial consensus
36	2026	82.1	467	20	AA43171	A. fumigatus phyta
37	2026	82.1	467	21	AA20528	Aspergillus fumiga
38	2026	82.1	467	21	AA69570	Mutant Aspergillus
39	1888	76.5	465	19	AAW84356	Aspergillus fumiga
40	1888	76.5	465	20	AA69905	A. fumigatus phyta
41	1878.5	76.1	440	21	AA69549	Aspergillus fumiga
42	1875.5	76.0	440	21	AA20507	Aspergillus fumiga
43	1875.5	76.0	440	21	AA69550	Aspergillus fumiga
44	1870.5	75.8	440	21	AA20508	Aspergillus fumiga
45	1870.5	75.8	440	21	AA69551	Aspergillus fumiga

ALIGNMENTS

RESULT 1

AA43170

ID AAY43170 standard; Protein; 467 AA.

XX AC AAY43170;

XX DT 06-JAN-2000 (first entry)

XX DE Consensus phytase-10-thermo(3)-Q50T-K91A protein sequence.

XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;

XX KW consensus sequence.

XX OS Synthetic.

XX PN WO9948380-A1.

XX PD 30-SEP-1999.

XX PF 22-MAR-1999; 99WO-DK00154.

XX PR 23-MAR-1998; 98DK-0000407.

XX PR 19-JUN-1998; 98DK-0000806.

XX PR 18-SEP-1998; 98DK-0001176.

XX PR 22-JAN-1999; 99DK-0000091.

XX PR 22-JAN-1999; 99DK-0000093.

XX PA (NOVO) NOVO-NORDISK AS.

XX PI Petersen S;

XX DR WPI; 1999-591030/50.

XX DR N-PSDB; AA231521.

PT Preparing animal feed using a thermostable phytase -

XX Example 3; Fig 10; 71pp; English.

XX This sequence represents the consensus phytase-10-thermo(3)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 20; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.6e-237;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPPFSLADE 60
 DB 1 mgvfvlslatlfgstgtalpgrnshscdvtvdggycfpeishlwgtyspffslade 60
 QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFLKTYN 120
 DB 61 saispdvpkgrvtvfqvlslrhgaryptssaskaysalieatoknatafkgyaflktn 120
 QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIETGQSA 180
 DB 121 ytlgaddltpfgeqomvnsikfyrykalkarkivpfirassgsdrviasaekfietgfsa 180
 QY 181 KLADPGANPHQASPVINVTIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvtipegagynntldhglctafeeseldgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDLTFHDEWIQYDYLQSLG 300
 DB 241 arleahlpvnltdedvvnldmcpdftvartsdatsqlspcdlthdewiqydyqlsgl 300
 QY 301 KYGAGNPLGPAQGVFVNLIELRLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvfvnlielrlthspvqdhstnthtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglngtkplsttsvesieetdgyaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlndrvvplhgcgvdkgkrdrdfveglsfarsggnweecfa 467

RESULT 2

AAB20527
 ID AAB20527 standard; Protein; 467 AA.

XX AAB20527;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 10 thermo 3 q50t, k91a protein SEQ ID NO:31.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytase degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

PN

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX N-PSDB; AAA73234.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Claim 4; Fig 8a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytase degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 21; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.6e-237;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTVDGGYQCFPEISHLWGTYSPPFSLADE 60
 DB 1 mgvfvlslatlfgstgtalpgrnshscdvtvdggycfpeishlwgtyspffslade 60
 QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFLKTYN 120
 DB 61 saispdvpkgrvtvfqvlslrhgaryptssaskaysalieatoknatafkgyaflktn 120
 QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIETGQSA 180
 DB 121 ytlgaddltpfgeqomvnsikfyrykalkarkivpfirassgsdrviasaekfietgfsa 180
 QY 181 KLADPGANPHQASPVINVTIPEGAGYNNLTDLHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladpganphqaspvinvtipegagynntldhglctafeeseldgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPFCDLTFHDEWIQYDYLQSLG 300
 DB 241 arleahlpvnltdedvvnldmcpdftvartsdatsqlspcdlthdewiqydyqlsgl 300
 QY 301 KYGAGNPLGPAQGVFVNLIELRLTHSPVQDHTSTNNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpaqgvfvnlielrlthspvqdhstnthtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdnmtwsiffalglngtkplsttsvesieetdgyaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGGNWEECEFA 467
 DB 421 lvrvlndrvvplhgcgvdkgkrdrdfveglsfarsggnweecfa 467

RESULT 3

AAV69569
 ID AAV69569 standard; Protein; 467 AA.
 XX
 AC AAV69569;
 XX
 DT 19-APR-2000 (first entry)
 XX
 DE Mutant phytase-10, phytase-10-thermo[3]-Q50T-K91A.
 XX
 KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 XX thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutein.
 XX
 OS Aspergillus terreus 9A1.
 OS Aspergillus terreus cbs16.46.
 OS Aspergillus niger var. awamori.
 OS Aspergillus niger str. NRRL3135.
 OS Aspergillus fumigatus ATCC13073.
 OS Aspergillus fumigatus ATCC32722.
 OS Aspergillus fumigatus ATCC58128.
 OS Aspergillus fumigatus ATCC26906.
 OS Aspergillus fumigatus ATCC32239.
 OS Emericella nidulans.
 OS Talaromyces thermophilus ATCC20186.
 OS Myceliophthora thermophila.
 OS Paxillus involutus NN005693.
 OS Trametes pubescens NN9343.
 OS Agrocybe pediades NN009289.
 OS Peniophora lycii NN006113.
 OS Thermomyces lanuginosa.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..26
 FT /note= "Phytase signal peptide from Aspergillus terreus
 FT cbs16.46"
 FT Protein 27..467
 FT /note= "Mature phytase-10-thermo[3]-Q50T-K91A"
 XX
 PN EP969089-A1.
 XX
 XX 05-JAN-2000.
 XX
 XX 23-JUN-1999; 99EP-0111949.
 XX
 XX 29-JUN-1998; 98EP-0111960.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA
 XX Bruggen R, Lehmann M, Wyss M;
 PI
 XX WPI; 2000-099429/09.
 DR
 XX
 XX New stabilised enzyme formulation, useful for feed compositions for
 PT monogastric animals
 PT
 XX
 XX Example 5; Fig 19; 101pp; English.
 PS
 XX
 XX The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra

CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-10
 CC consensus sequence, phytase-10-thermo[3]-Q50T-K91A, which has a
 CC temperature optimum and melting point 4 degrees Celsius higher than that
 CC of phytase-10 (AAV69566). Its specific activity with phytate as a
 CC substrate is also strongly increased.
 XX
 XX Sequence 467 AA;

Query Match 100.0%; Score 2468; DB 21; Length 467;
 Best Local Similarity 100.0%; Pred. No. 2.6e-237;
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVFVLLSIATLFGSTGCTALGPRGNHSCDTVDGGYQCFFPEISHLMWTYSPPFSLADE 60
 DB 1 mgvfvllsiatlfgstgctalgrgnshscdtvdggyqcfeishlwtyspffslade 60
 QY 61 SAISPDVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFLKTYN 120
 DB 61 saispdvpgkcrvtfvqlsrhgaryptssaskaysalieleaigknatarfgyaflkty 120
 QY 121 YTLGADLTFFGEEQOMVNSGIKFRYKALARKIVPFIKASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltfgeeqomvnsigikfrykalarakiwpfirasgsdrviasaeekfieqf 180
 QY 181 KLADFCANPHQASPVINVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 DB 181 kladfcanphqaspvinipegagynntldhglctafeeselgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSDFCDLTFHDEWIOYDIQSLG 300
 DB 241 arleahlpgvnltdvnlmdmcpdftvartsdatsdqslpfdcltfhdewidydyiqslg 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 DB 301 kyygygagnplgpagvgfvneliarlthspvgdhtstnhtldsnpatplnatlyadfs 360
 QY 361 HDNTWVSIFPFFALGLYNGTKPLSTTSVESIEETDGSASWTVPFAARAYVEMMQCEAKEP 420
 DB 361 hdntwvsiffalglyngtkplsttsvesieetdgsaswtvpfaarayvemmqceakep 420
 QY 421 LVRVLVNDVRVPLHGGVDKLGRCRKRDDPVEGLSFARSGGNWEECFA 467
 DB 421 lvrvlvndvrvpilhggvdkgkrckrddpveglstarsggnweecfa 467

RESULT 4

AAAB20534
 ID AAB20534 standard; Protein; 467 AA.
 XX
 AC AAB20534;
 XX
 DT 05-DEC-2000 (first entry)
 XX
 DE Consensus phytase 10 thermo 5 Q50T K91A protein SEQ ID NO:97.
 XX
 KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX
 OS Synthetic.
 XX
 XX WO200043503-A1.
 PN

XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-DK00025.
XX PR 22-JAN-1999; .99DK-0000092.
XX PR 21-SEP-1999; 99DK-0001340.
XX PA (NOVO) NOVO NORDISK AS.
XX PI Lehmann M;
XX DR WPI; 2000-491161/43.
XX DR N-PSDB; AAA73293.
XX PT Novel phytases with improved properties such as temperature stability,
XX PT pH stability and substrate specificity, for use in pharmaceuticals and
XX PT compound foods and feeds -
XX PS Disclosure; Fig 25a-c; 240pp; English.
XX CC The present invention describes improved phytases, preferably with
XX CC increased thermostability, and methods for producing them. The methods
XX CC can be used for producing phytases with improved properties e.g.
XX CC temperature stability, pH stability, pH profile, temperature profile,
XX CC specific activity, substrate specificity, substrate cleavage pattern,
XX CC substrate binding, position specificity, the velocity and level of
XX CC release of phosphate from corn, reaction rate, phytate degradation rate,
XX CC and end level of released phosphate. The phytases can be used to produce
XX CC pharmaceutical compositions or compound food or feeds. The feed can be
XX CC used to reduce levels of phytate in animal manure, by converting it
XX CC into lower inositol phosphates and/or inositol and inorganic phosphate.
XX CC The present sequence represents a phytase sequence from the present
XX CC invention.
XX SQ Sequence 467 AA;
Query Match 99.5%; Score 2456; DB 21; Length 467;
Best Local Similarity 99.6%; Pred. No. 4.1e-236;
Matches 465; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MGFFVLLSIATLFGSTGTALGRGNHSCDVTGQYQCFPEISHLWGTYSPPFFSLADE 60
DB 1 mgffvllsiatlfgstgtalgrgnshscdvtgqyqcfpeishlwgtysppffslade 60
QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNAFAKGYAFLKTYN 120
DB 61 saispdpvkgcrvtfvqlsrhgaryptssaskaysalieatqknafakgyaflkty 120
QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFOSA 180
DB 121 ytlgaddltpfgqomvnschkfryrykalkarkivpfiirassgdrviasaekfiqf 180
QY 181 KIADPGANPHQASPVNIIIEGAGYNNITLHGICTAFEESELGDDVEANTVAFAPPIR 240
DB 181 kiadpganphqaspvniiliegagynntldhgictafeeselgddveantvafappi 240
QY 241 ARLEAHLPGVNLTDDEYVNLMDMCPEDTAVRTSDATQSPFCDLTFHDEWIOYDYLQSLG 300
DB 241 arleahlpgvnltdedvynlmdmcpedfvartsdatsqlspfdlthdewioydyqls 300
QY 301 KYGYGAGNPLGPAGCGVFNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
DB 301 kygygagnplgpagcgvfneliarlthspvqdhtstnhtldsnpatfplnatlyads 360
QY 361 HNTWVSVIFFALGYNCKPKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCEAKEP 420
DB 361 hntwsvsiffalgyngtkplsttsvesieetdgyasaswtvpfaarayvemmqceake 420
QY 421 LVRLVNDRVVPLHCGGVDKLGRCKRDDDFVEGLSFARSGGNWEECF 467
DB 421 lvrlvndrvvplhgcgvdklgrckrdddfveglsfarsggnweecfa 467

RESULT 5
AAB20533
ID AAB20533 standard; Protein; 467 AA.
XX AC AAB20533;
XX DT 05-DEC-2000 (first entry)
XX DE Consensus phytase 10 thermo 5 Q50T protein SEQ ID NO:95.
XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
XX KW temperature stability; pH profile; temperature profile; reaction rate;
XX KW specific activity; substrate specificity; substrate cleavage pattern;
XX KW substrate binding; position specificity; phytate degradation rate;
XX KW food; feed; phytate; manure.
XX OS Synthetic.
XX PN WO2000043503-A1.
XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-DK00025.
XX PR 22-JAN-1999; 99DK-0000092.
XX PR 21-SEP-1999; 99DK-0001340.
XX PA (NOVO) NOVO NORDISK AS.
XX PI Lehmann M;
XX DR WPI; 2000-491161/43.
XX DR N-PSDB; AAA73292.
XX PT Novel phytases with improved properties such as temperature stability,
XX PT pH stability and substrate specificity, for use in pharmaceuticals and
XX PT compound foods and feeds -
XX PS Disclosure; Fig 24a-c; 240pp; English.
XX CC The present invention describes improved phytases, preferably with
XX CC increased thermostability, and methods for producing them. The methods
XX CC can be used for producing phytases with improved properties e.g.
XX CC temperature stability, pH stability, pH profile, temperature profile,
XX CC specific activity, substrate specificity, substrate cleavage pattern,
XX CC substrate binding, position specificity, the velocity and level of
XX CC release of phosphate from corn, reaction rate, phytate degradation rate,
XX CC and end level of released phosphate. The phytases can be used to produce
XX CC pharmaceutical compositions or compound food or feeds. The feed can be
XX CC used to reduce levels of phytate in animal manure, by converting it
XX CC into lower inositol phosphates and/or inositol and inorganic phosphate.
XX CC The present sequence represents a phytase sequence from the present
XX CC invention.
XX SQ Sequence 467 AA;
Query Match 99.3%; Score 2451; DB 21; Length 467;
Best Local Similarity 99.4%; Pred. No. 1.3e-235;
Matches 464; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MGFFVLLSIATLFGSTGTALGRGNHSCDVTGQYQCFPEISHLWGTYSPPFFSLADE 60
DB 1 mgffvllsiatlfgstgtalgrgnshscdvtgqyqcfpeishlwgtysppffslade 60
QY 61 SAISPDPVKGCRVTFVQVLSRHGARYPTSSASKAYSALIEATQKNAFAKGYAFLKTYN 120
DB 61 saispdpvkgcrvtfvqlsrhgaryptssaskaysalieatqknafakgyaflkty 120
QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIIRASGSDRVIASAEKFIQFOSA 180
DB 121 ytlgaddltpfgqomvnschkfryrykalkarkivpfiirassgdrviasaekfiqf 180

Db 121 ytlgaddltpfgeqgnvnsigikfyrtrykalarakivpfirasgdrviaseakfiegfsa 180
 QY 181 KLADGANPHQASPVINVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 Db 181 kladganphqaspviniipegagynntldhglctafeestlgddveanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSPFCDLFTHEWIOYDYLQSLG 300
 Db 241 arleahlpvnltdedvnlmdmcpdftvartsdatsqlspfcldfhdewigydyqlslg 300
 QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 Db 301 kygygagnp1gpadgvgfvneliarlthspvgdhtstnhtldsnpatplnatlyadfs 360
 QY 361 HDNTMVSIFPFGALGNGTKPLSTTSVESTEEDGYASWTVPFAARAYVEMMOCEAKEP 420
 Db 361 hdntmvsiffalglngtkplsttsvesieetdgyaswcvpfaarayvemmocaeakep 420
 QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467
 Db 421 lvrvlndrvvplhgcavdklgrckrddfvglslfarsggnweecfa 467

RESULT 6

AAB20524
 ID AAB20524 standard; Protein: 467 AA.

XX AC AAB20524;

DT 05-DEC-2000 (first entry)

DE Consensus phytase 10 SEQ ID NO:26.

KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

PN WO200043503-A1.

PD 27-JUL-2000.

PF 21-JAN-2000; 2000WO-DK000025.

PR 22-JAN-1999; 99DR-00000092.

PR 21-SEP-1999; 99DR-0001340.

PA (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI: 2000-491161/43.

XX N-PSDB; AAA73232.

PT Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 1; Fig 5a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g. file,
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it

CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

SQ Sequence 467 AA;

Query Match 99.2%; Score 2448; DB 21; Length 467;

Best Local Similarity 98.9%; Pred. No. 2.6e-235;

Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGVEVLLSIATLFGSTGTALGPRGNHSCDVGQYQCFPEISHLWGTYSFFSLADE 60

Db 1 mgvfvllsiatlfgstgtalgrgnshscdvdggyqcfpeishlwgtysffslade 60

QY 61 SAISPDVPRGCRVTFVQVLSRHGARYPTSSAKAYASALIEAQKNATAFKGVAFUKTYN 120

Db 61 saispdvprkgcrvtfvqlsrhgaryptsakkskysallieaigknatafkgvafukty 120

QY 121 YTLGADDLTPFGEQGMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180

Db 121 ytlgaddltpfgeqgmvnsgikfyrtrykalarakivpfirasgdrviaseakfiegfsa 180

QY 181 KLADGANPHQASPVINVIPEGAGYNNITLDHGLCTAFEESELGDDVEANFTAVFAPP 240

Db 181 kladganphqaspviniipegagynntldhglctafeestlgddveanftavfappir 240

QY 241 ARLEAHLPGVNLTDVNLMDMCPDFTVARTSDATQLSPFCDLFTHEWIOYDYLQSLG 300

Db 241 arleahlpvnltdedvnlmdmcpdftvartsdatsqlspfcldfhdewigydyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

Db 301 kygygagnp1gpadgvgfvneliarlthspvgdhtstnhtldsnpatplnatlyadfs 360

QY 361 HDNTMVSIFPFGALGNGTKPLSTTSVESTEEDGYASWTVPFAARAYVEMMOCEAKEP 420

Db 361 hdntmvsiffalglngtkplsttsvesieetdgyaswcvpfaarayvemmocaeakep 420

QY 421 LVRVLNDRVPLHGGVDKLGRCRDRDFVEGLSFARSGGNWEECFA 467

Db 421 lvrvlndrvvplhgcavdklgrckrddfvglslfarsggnweecfa 467

RESULT 7

AAAY69566 standard; protein: 467 AA.

XX AC AAAY69566;

XX 19-APR-2000 (first entry)

DE Phytase-10, a consensus phytase.

XX Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;

XX thermotable; animal feed; monogastric animal; phytate phosphorus;

XX phosphate availability; consensus; phytase-10.

OS Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS Aspergillus niger var. awamori.

OS Aspergillus niger str. NRRL3135.

OS Aspergillus fumigatus ATCC13073.

OS Aspergillus fumigatus ATCC32722.

OS Aspergillus fumigatus ATCC58128.

OS Aspergillus fumigatus ATCC26906.

OS Emericella nidulans.

OS Talaromyces thermophilus ATCC20186.

OS Myceliophthora thermophila.

OS Paxillus involutus NN005693.

OS Trametes pubescens NN9343.

OS Agrocyste pediales NN009289.

121	YTLGADDLTPFGEOQMVNSG	IKFYRRYKALARKIVPIR	ASGSDRVITASAEKFL	EGQSA	180		
121	YTLGADDLTPFGEOQMVNSG	IKFYRRYKALARKIVPIR	ASGSDRVITASAEKFL	EGQSA	180		
121	YTLGADDLTPFGEOQMVNSG	IKFYRRYKALARKIVPIR	ASGSDRVITASAEKFL	EGQSA	180		
161	KLADPGANPHQASPVINVI	PEGAGYNNTLDHGLCTA	FESELGDDVEANFTAV	FAPPIR	240		
161	KLADPGANPHQASPVINVI	PEGAGYNNTLDHGLCTA	FESELGDDVEANFTAV	FAPPIR	240		
161	KLADPGANPHQASPVINVI	PEGAGYNNTLDHGLCTA	FESELGDDVEANFTAV	FAPPIR	240		
241	ARLEAHLPGVNLTD	EDVNVNMDMCP	FFTVARTS	DATQLSP	PCDLFTHDEWIDYDLSLG	300	
241	ARLEAHLPGVNLTD	EDVNVNMDMCP	FFTVARTS	DATQLSP	PCDLFTHDEWIDYDLSLG	300	
301	KYGYGAGNPLCPAGGV	GVFNELIARLTHSP	QVQDHTSTNNHTL	DSNPAT	EPLNATLYADFS	360	
301	KYGYGAGNPLCPAGGV	GVFNELIARLTHSP	QVQDHTSTNNHTL	DSNPAT	EPLNATLYADFS	360	
361	HDNTMVSIF	FALLGYNGTKPLSTTS	VESIEETD	GYSAWTV	PFARAYVEMQCEA	KEP	420
361	HDNTMVSIF	FALLGYNGTKPLSTTS	VESIEETD	GYSAWTV	PFARAYVEMQCEA	KEP	420
421	LVRVLVNDVRVPLHCG	GVYDKLGRCKR	DDFVEGL	SPARSGGNWEE	CFA	467	
421	LVRVLVNDVRVPLHCG	GVYDKLGRCKR	DDFVEGL	SPARSGGNWEE	CFA	467	
421	LVRVLVNDVRVPLHCG	GVYDKLGRCKR	DDFVEGL	SPARSGGNWEE	CFA	467	
RESULT	8						
AAB20532							
ID	AAB20532	standard; Protein: 467	AA.				
XX							
AC	AAB20532;						
XX							
DT	05-DEC-2000	(first entry)					
DE	Consensus	phytase 3 thermo 11 Q50T K91A	protein SEQ ID NO:93.				
XX	Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;						
XX	temperature stability; pH profile; temperature profile; reaction rate;						
XX	specific activity; substrate specificity; substrate cleavage pattern;						
XX	substrate binding; position specificity; phytate degradation rate;						
XX	food; feed; phytate; manure.						
OS	Synthetic.						
XX							
PN	WO200043503-A1.						
PD	27-JUL-2000.						
EF	21-JAN-2000; 2000WO-DK00025.						
XX							
PR	22-JAN-1999; 99DK-0000092.						
PR	21-SEP-1999; 99DK-0001340.						
XX							
PA	(NOVO) NOVO NORDISK AS.						
XX							
PI	Lehmann M;						
DR	WPI; 2000-491161/43.						
DR	N-PSDB; AAA73291.						
XX							
PT	Novel phytases with improved properties such as temperature stability						
PT	pH stability and substrate specificity, for use in pharmaceuticals and						
PT	compound foods and feeds -						
XX	Disclosure; Fig 23a-c; 240pp; English.						
XX							
CC	The present invention describes improved phytases, preferably with						
CC	increased thermostability, and methods for producing them. The method						
CC	can be used for producing phytases with improved properties e.g.						
CC	temperature stability, pH stability, pH profile, temperature profile,						
CC	specific activity, substrate specificity, substrate cleavage pattern,						
CC	substrate binding, position specificity, the velocity and level of						
CC	release of phosphate from corn, reaction rate, phytate degradation ra						
CC	and end level of released phosphate. The phytases can be used to prod						

OS	Penicillium laticellulosum	XX
OS	Thermomyces lanuginosa.	XX
OS	Synthetic.	XX
XX	Key	XX
XX	Location/Qualifiers	XX
XX	1..26	XX
XX	Peptide	XX
XX	/note= "Phytase signal peptide from Aspergillus terreus cbs16.46"	XX
XX	27..467	XX
XX	Protein	XX
XX	/note= "Mature phytase-10 consensus"	XX
XX	23-JUN-1999; 99EP-0111949.	XX
XX	29-JUN-1998; 98EP-0111960.	XX
XX	(HOFF) HOFFMANN LA ROCHE & CO AG.F.	XX
XX	Brugger R, Lehmann M, Wyss M;	XX
XX	WPI: 2000-099429/09.	XX
XX	N-PSDB; AAZ59642.	XX
XX	New stabilised enzyme formulation, useful for feed compositions for monogastric animals -	XX
XX	Example 4; Fig 17; i01pp; English.	XX
XX	The invention relates to a novel stabilised dry or liquid enzyme formulation, comprising phytase (myo-inositol hexakisphosphate phosphohydrolase) and one or more stabilising agents including xylitol or ribitol; polyethylene glycols with a molecular weight of 600 to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic, glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.	XX
XX	The stabilised phytase formulation is used in a method for preparing a feed composition for monogastric animals (e.g., pigs, poultry) and provides a monogastric animal with its dietary requirements of phosphorus. Although a large amount of phosphate is present in animal feed in the form of phytate phosphorus, monogastric animals are unable to utilise this form of phosphate, resulting in the addition of extra phosphate to the feed of such animals. Phytase enhances the nutritional value of plant material without the need for adding additional phosphate to the feed. The level of phosphate pollution in the environment is reduced by adding phytase to animal feed, as the animal can make use of the inorganic phosphate liberated from phytate phosphorus using the enzyme. The phytase formulation of the invention has an improved thermostability and can therefore remain stable during long-term storage and can withstand feed processing methods such as extrusion, expansion and pelleting. The present sequence represents a consensus sequence, designated Phytase-10, which was derived from the mature phytase sequences from a variety of fungi (AAV69544-f69546, AAV69548-f69556, AAV69564) and the Basidiomycetes phytase consensus AAV69563 and additionally contains the Aspergillus terreus cbs16.46 signal peptide at the N-terminus.	XX
XX	Sequence	XX
XX	467 AA;	XX
XX	Query Match	XX
XX	Best Local Similarity 99.2%; Score 2448; DB 21; Length 467;	XX
XX	Matches 462; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX
XX	1 MGVFVLLSLATLFGSTSTALGPRGNSHSDTVGDGQCQCFPEISHLWGYSPFFSLADE 60	XX
XX	1 mgvfvlslatlfgststalgprgnshsdctvvgqcfpeishlwgyspffslade 60	XX
XX	61 SAISPDVPGRCVTFVQVLSRHRGARYPTSSASKAYSALIEAIOKNATFAFGKYAFKLTYN 120	XX
XX	61 saispdvpgrcvtfvqlsrhgaryptsskysalialieaigknaatfagkyafikltn 120	XX

CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.9%; Score 2367; DB 21; Length 467;
 Best Local Similarity 95.7%; Pred. No. 3.1e-227;
 Matches 447; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGCGYCFPEISHLWGTSPFFSLADE 60
 Db 1 mgvfvlslatlfgstgaltgprgnshscdtvdcgycfpeishlwgtspffslade 60

QY 61 SAISDPVPGRCVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAPKGYAFILKTYN 120
 Db 61 saispdvpkdcrtvfvqlsrhgaryptssaskaysalieaoknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYRKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180
 Db 121 ytlgaddltpfgengmvsngikfrrykalarkivpfiragsdrviaseakfiqfqs 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpgsqhqsapvlnvipegsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWIQYDLOSLG 300
 Db 241 arleahlpvnltddevvnlmdcpfdvtartsdatselgddveanftalfapair 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360
 Db 301 kygygagnpplgagvgvfanellrthspvqdhstntldsnpatplnatlyadfs 360

QY 361 HNTMVSIFFAALGNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hntmvsiffalglngtkplsttsvesieetdgyasaswtvpfaarayvemmqqakep 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFEGLSFARSGGNWEECA 467
 Db 421 lvrvlndrvplhgcavdklgrckdrdfveglsfarsggnwaeca 467

RESULT 9

AAV43169
 ID AAV43169 standard; Protein; 467 AA.

XX AC AAY43169;

XX DT 06-JAN-2000 (first entry)

XX DE Consensus phytase-1-thermo(8)-Q50T-K91A protein sequence.

XX KW Phytase; animal feed preparation; thermostable phytase; transgenic plant;
 consensus sequence.

XX OS Synthetic.

XX PN WO9948380-A1.

XX PD 30-SEP-1999.

XX PF 22-MAR-1999; 99WO-DK00154.

XX PR 23-MAR-1998; 98DK-0000407.

XX PR 19-JUN-1998; 98DK-0000806.

XX PR 18-SEP-1998; 98DK-0001176.

XX PR 22-JAN-1999; 99DK-0000091.

XX PR 22-JAN-1999; 99DK-0000093.

PA (NOVO) NOVO-NORDISK AS.

XX PI Petersen S;

XX WI 1999-591030/50.

XX DR N-PSDB; AA231520.

XX PT Preparing animal feed using a thermostable phytase

XX Example 3; Fig 9; 7lpp; English.

CC This sequence represents the consensus phytase-1-thermo(8)-Q50T-K91A.
 CC The invention relates to a process for preparing animal feed by
 CC agglomerating feed ingredients with a thermostable phytase, which is
 CC added before or during agglomeration. The thermostable phytase is useful
 CC for expression in transgenic plants. These plants are useful in the
 CC preparation of animal feed itself. The thermostable phytase allows animal
 CC feed to be produced more efficiently, in addition to improved
 CC phytase-expressing transgenic plants. These plants provide a feed
 CC ingredient and a feed additive (phytase) simultaneously.

XX Sequence 467 AA;

Query Match 95.7%; Score 2362; DB 20; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGVFVLLSIATLFGSTGALPGRNHSCDVTGCGYCFPEISHLWGTSPFFSLADE 60

Db 1 mgvfvlslatlfgstgaltgprgnshscdtvdcgycfpeishlwgtspffslade 60

QY 61 SAISDPVPGRCVTFVQVLSRHGARYPTSSASKAYSALIEAOKNATAPKGYAFILKTYN 120

Db 61 saispdvpkdcrtvfvqlsrhgaryptssaskaysalieaoknatafkgyafilktyn 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYRKALARKIVPFIIRASGSDRVIAAEKFIQFQSA 180

Db 121 ytlgaddltpfgengmvsngikfrrykalarkivpfiragsdrviaseakfiqfqs 180

QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240

Db 181 kladpgsqhqsapvlnvipegsgynntldhglctafedstlgddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDATQSPFCDLTFHDEWIQYDLOSLG 300

Db 241 arleahlpvnltddevvnlmdcpfdvtartsdatselgddveanftalfapair 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATPLNATLYADFS 360

Db 301 kygygagnpplgagvgvfanellrthspvqdhstntldsnpatplnatlyadfs 360

QY 361 HNTMVSIFFAALGNGTKPLSTTSVESIEETDGYASASWTVPFAARAYVEMMQCEAKEP 420

Db 361 hntmvsiffalglngtkplsttsvesieetdgyasaswtvpfaarayvemmqqakep 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFEGLSFARSGGNWEECA 467

Db 421 lvrvlndrvplhgcavdklgrckdrdfveglsfarsggnwaeca 467

RESULT 10

AAAB20526

ID AAB20526 standard; Protein; 467 AA.

XX AC AAB20526;

XX DT 05-DEC-2000 (first entry)

XX DE Consensus phytase 1 thermo 8 q50t, k91a protein SEQ ID NO:29.

XX KW Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 temperature stability; pH profile; temperature profile; reaction rate;

KW specific activity; substrate specificity; phytate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.
 XX Synthetic.

OS WO200043503-A1.

XX 27-JUL-2000.

PF 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73233.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Claim 5; Fig 7a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.7%; Score 2362; DB 21; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCPEISHLWGTSPFFSLADE 60
 DB 1 mgvfvllsiatlfgstgtalgpgrnshscdtpdggycpfeishlwgtspffslade 60
 QY 61 SAISDPVPGKCRVTFVQLSRGARYPTSSASKAYSALIEAOKNATAFKGYAFLKTYN 120
 DB 61 saisdvpdpcrcvtfvqlsrhgarpytssaskaysalieaoknatafkgyafktyln 120
 QY 121 YTLGADDLTPFGEOQMNWNGIKFYRRYKALARKIYPTFRASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgengmnwngikfrrykalarkiyptfrasgsdrviasaekfiqfqsaa 180
 QY 181 KLADPGANPHQASPVINVIPEGAGYNNLTDLHGLCTAFEESELGDDVDFANFTAVFAPPIR 240
 DB 181 kladpgsqhqpaspvlnvipegagynntldhgtctafedseelgddveanfalfapair 240
 QY 241 ARLEAHLPGVNLTDENVLMDCPFDTVARTSDATLSPFCDLTHDEWQYDYLQSLG 300
 DB 241 arleahlpvnltdenvylmcpfdtvtartsdatslspfcldthdewqydyqlsgl 300
 QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPFNATLYADF 360
 DB 301 kyygygagnpapgvgvfganeliarlthspvqdhststnhtldsnpatfpfnatlyadfs 360

QY 361 HDNTMWSIFFALGLYNGTKPLSTTSVESIEETDGYSSASWTVPFAARAYVEMMQCAKEP 420
 DB 361 hdnmtwsiffalglyngtkplsttsvesieetdgyssaswtvpfaarayvemmqcdaekep 420
 QY 421 LVRVLYNDRVVPVPHGCGVDKLGRCRDRDDFVEGLSFARSGNWECEFA 467
 DB 421 lvrvlyndrvvpvphgcavdklgrckrddfveglsfarsgnwaeacfa 467

RESULT 11

AAB20531

ID AAB20531 standard; Protein; 467 AA.

XX AAB20531;

XX 05-DEC-2000 (first entry)

XX Consensus phytase 3 thermo 11 Q50T protein SEQ ID NO:91.

XX Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

OS Synthetic.

XX WO2000043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

PR 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

DR N-PSDB; AAA73290.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

PS Disclosure; Fig 2a-c; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 467 AA;

Query Match 95.7%; Score 2362; DB 21; Length 467;
 Best Local Similarity 95.5%; Pred. No. 9.7e-227;
 Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGFFVLLSIATLFGSTGTALGPRGNSHSCDTPDGGYQCPEISHLWGTSPFFSLADE 60
 DB 1 mgvfvllsiatlfgstgtalgpgrnshscdtpdggycpfeishlwgtspffslade 60

QY 61 SAISPDVPGKCRVTFVQVLSRHRGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120
 DB 61 saispdvpgkcrvtfvqvlshrhrgaryptssaskaysalieleaiqknatafkgyafilkty 120
 QY 121 YTLGADDLTPFGQQQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
 DB 121 ytlgaddltpfgqqqmvnsgikfrryrykalarkivpfirasgsdrviasaekfiqfqs 180
 QY 181 KLADPGANPQASPVINVIPEGAGVNNLTLDHGLCTAPESLGDVDEANFTAVFAPPIR 240
 DB 181 kladpganpqaspvinvipegagvnnltdhglctapeslgdvdeanftavfappir 240
 QY 241 ARLEAHLPGVNLTDVNNLMDCPFDVARTSDATQISPFCDLTHDEWIDYDYLQSLG 300
 DB 241 arleahlpgvnltdvnnlmdcpfdvartsdatspfcldthdewidydyqlsg 300
 QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHTSNHTLDSNPATFPLNATLYADFS 360
 DB 301 kygygagnplgpagvfneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360
 QY 361 HDNTMWSIFFALGLXNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420
 DB 361 hdntmwsiffalglxngtkplsttsvesieetdgyaswtvpfaarayvemmqcaekep 420
 QY 421 LYRVLVNDRVPLHGGGVYDKLGRCKRDDFVEGLSFARSGNNEECPA 467
 DB 421 lyrvlvndrvplhgggvdydklgrckrddfveglsfarsgnneecpa 467

RESULT 12

AAV69568

ID AAV69568 standard; protein; 467 AA.

XX AC AAV69568;

DT 19-APR-2000 (first entry)

DE Mutant phytase-1, phytase-1-thermo[8]-Q50T-K91A.

KW Phytase; myo-inositol hexakisphosphate phosphohydrolase; stabilisation;
 KW thermostable; animal feed; monogastric animal; phytate phosphorus;
 KW phosphate availability; consensus; mutant; mutcin.

OS Aspergillus terreus 9A1.

OS Aspergillus terreus cbs16.46.

OS Aspergillus niger var. awamori.

OS Aspergillus niger t213.

OS Aspergillus niger str. NRRL3135.

OS Aspergillus fumigatus ATCC13073.

OS Aspergillus fumigatus ATCC32722.

OS Aspergillus fumigatus ATCC58128.

OS Aspergillus fumigatus ATCC26906.

OS Emericella nidulans.

OS Talaromyces thermophilus ATCC20186.

OS Myceliophthora thermophila.

OS Synthetic.

FH Key Location/Qualifiers

FT Peptide 1..26

FT /note= "Phytase signal peptide from Aspergillus terreus

FT /note= cbs16.46"

FT Protein 27..467

FT /note= "Mature phytase-1-thermo[8]-Q50T-K91A"

XX EP969089-A1.

XX 05-JAN-2000.

XX 23-JUN-1999;

XX 99EP-0111949.

XX 29-JUN-1998;

XX 98EP-0111960.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX Brugger R, Lehmann M, Wyss M;

XX WPI: 2000-099429/09.

XX N-PSDB; AAZ59715.

XX New stabilized enzyme formulation, useful for feed compositions for

XX monogastric animals -

XX Example 5; Fig 19; 101pp; English.

XX The invention relates to a novel stabilised dry or liquid enzyme
 CC formulation, comprising phytase (myo-inositol hexakisphosphate
 CC phosphohydrolase) and one or more stabilising agents including
 CC xylitol or ribitol; polyethylene glycols with a molecular weight of 600
 CC to 4000 Da, preferably 1000 to 3350 Da; the disodium salts of malonic,
 CC glutaric and succinic acid; carboxymethylcellulose; and sodium alginate.
 CC The stabilised phytase formulation is used in a method for preparing a
 CC feed composition for monogastric animals (e.g., pigs, poultry) and
 CC provides a monogastric animal with its dietary requirements of
 CC phosphorus. Although a large amount of phosphate is present in animal
 CC feed in the form of phytate phosphorus, monogastric animals are unable
 CC to utilise this form of phosphate, resulting in the addition of extra
 CC phosphate to the feed of such animals. Phytase enhances the nutritional
 CC value of plant material without the need for adding additional phosphate
 CC to the feed. The level of phosphate pollution in the environment is
 CC reduced by adding phytase to animal feed, as the animal can make use of
 CC the inorganic phosphate liberated from phytate phosphorus using the
 CC enzyme. The phytase formulation of the invention has an improved
 CC thermostability and can therefore remain stable during long-term storage
 CC and can withstand feed processing methods such as extrusion, expansion
 CC and pelleting. The present sequence represents a mutant phytase-1
 CC consensus sequence, phytase-1-thermo[8]-Q50T-K91A, which has a
 CC temperature optimum and melting point 7 degrees Celsius higher than that
 CC of phytase-1 (AAV69558).

XX Sequence 467 AA;

Query Match

Best Local Similarity 95.7%; Score 2362; DB 21; Length 467;

Matches 446; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLFGSTGTALGPRGNHSCDTVDGGYQCFFPEISHLWGTSPFFSLADE 60

DB 1 mgvfvlslatlfgstgtalgrgnshscdtvdggyqcfeishlwgtspysfslade 60

QY 61 SAISPDVPGKCRVTFVQVLSRHRGARYPTSSASKAYSALIEAIQKNATAFKGYAFILKTYN 120

DB 61 saispdvpgkcrvtfvqvlshrhrgaryptssaskaysalieleaiqknatafkgyafilkty 120

QY 121 YTLGADDLTPFGQQQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180

DB 121 ytlgaddltpfgqqqmvnsgikfrryrykalarkivpfirasgsdrviasaekfiqfqs 180

QY 181 KLADPGANPQASPVINVIPEGAGVNNLTLDHGLCTAPESLGDVDEANFTAVFAPPIR 240

DB 181 kladpganpqaspvinvipegagvnnltdhglctapeslgdvdeanftavfappir 240

QY 241 ARLEAHLPGVNLTDVNNLMDCPFDVARTSDATQISPFCDLTHDEWIDYDYLQSLG 300

DB 241 arleahlpgvnltdvnnlmdcpfdvartsdatspfcldthdewidydyqlsg 300

QY 301 KYGYGAGNPLGPAQGVFNELIARLTHSPVODHTSNHTLDSNPATFPLNATLYADFS 360

DB 301 kygygagnplgpagvfneliarlthspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNTMWSIFFALGLXNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAKEP 420

DB 361 hdntmwsiffalglxngtkplsttsvesieetdgyaswtvpfaarayvemmqcaekep 420

QY 421 LVRVLNDRVPLHCGGVYDKLGRCKRDDVFGLSFARSGGNWEECF 467
 DB 421 lvrvlndrvplhgcavdklgrckrddvfglsfarsggnwaecfa 467

RESULT 13

AAAB20523
 ID AAB20523 standard; Protein; 441 AA.

AC AAB20523;

XX 05-DEC-2000 (first entry)

DT Consensus phytase 10 (Fcp10) SEQ ID NO:24.

DE Phytase; mutant; thermostability; mutation; mutagenesis; pH stability;
 KW temperature stability; pH profile; temperature profile; reaction rate;
 KW specific activity; substrate specificity; substrate cleavage pattern;
 KW substrate binding; position specificity; phytate degradation rate;
 KW food; feed; phytate; manure.

XX Synthetic.

XX WO200043503-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-DK00025.

XX 22-JAN-1999; 99DK-0000092.

XX 21-SEP-1999; 99DK-0001340.

XX (NOVO) NOVO NORDISK AS.

XX Lehmann M;

XX WPI; 2000-491161/43.

XX Novel phytases with improved properties such as temperature stability,
 PT pH stability and substrate specificity, for use in pharmaceuticals and
 PT compound foods and feeds -

XX Example 2; Fig 4a-d; 240pp; English.

XX The present invention describes improved phytases, preferably with
 CC increased thermostability, and methods for producing them. The methods
 CC can be used for producing phytases with improved properties e.g.
 CC temperature stability, pH stability, pH profile, temperature profile,
 CC specific activity, substrate specificity, substrate cleavage pattern,
 CC substrate binding, position specificity, the velocity and level of
 CC release of phosphate from corn, reaction rate, phytate degradation rate,
 CC and end level of released phosphate. The phytases can be used to produce
 CC pharmaceutical compositions or compound food or feeds. The feed can be
 CC used to reduce levels of phytate in animal manure, by converting it
 CC into lower inositol phosphates and/or inositol and inorganic phosphate.
 CC The present sequence represents a phytase sequence from the present
 CC invention.

XX Sequence 441 AA;

Query Match 94.1%; Score 2322; DB 21; Length 441;
 Best Local Similarity 98.9%; Pred. No. 8.6e-223;
 Matches 436; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 27 NSHSCDVGCGYCFPIISHLWGTYSPPFFSLADESAISDPVPGKCRVTFVQLSRGARY 86
 DB 1 nshscdvdyggyqcfpeishlwqyspffsladesaisdpvpgkcrvtvqvlrsrgary 60

QY 87 PTSSASRAYSALTEATOKNATKGYAFKTYNTLIGADDLPPFGQQMVNSGKIFYRR 146
 DB 61 ptsskskysallaeaknatafkgyafktyntlgadddltpfgeqmvnsqikfyr 120

QY 147 YKALARKIVPFRASGSDRVIASAEKFIQFOSAKLADPGANPHOASPVINVIPEGAGY 206
 DB 121 ykalarlkivpfrasgsdrviasaekflegfsakladpganphaspsvinvilpegagy 180
 QY 207 NNTLDHGLCTAFEESELGDDEANFTAVFAPPTRARLEAHLPGVNLTDEDVYNLMDMCPF 266
 DB 181 nntldhglctafeeseldgdeanftavfapptrarleahlpgvnltdeedvynlmdmcpf 240
 QY 267 DTVARTSDATQSLSPFCDLTFHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIAR 326
 DB 241 dtvartsdattqlspfcldtfhdewiodylqslgkygygagnpigpaqgvgfvnelliar 300
 QY 327 LTHSPVQDHTSTNHTLDSNPATPLNATLVADFSDHNTWVSIFFALGLYNGTKPLSTTSV 386
 DB 301 lthspvqdhtstnhtldsnpatplnatlvadfshdntwvsiffalglyngtkplsttsv 360
 QY 387 ESIEETDGYASWTVPFAARAYVEMMOCAEKEPLVRVLNDRVPLHCGGVYDKLGRCKR 446
 DB 361 esieetdgyaaswtvpfaarayvemmqcaakeplvrvlndrvplhgcgvdklgrckr 420
 QY 447 DDFVEGLSFARSGGNWEECF 467
 DB 421 ddfveglsfarsggnweecfa 441

RESULT 14

AAW93382
 ID AAW93382 standard; Protein; 467 AA.

XX AAW93382;

XX 11-JUN-1999 (first entry)

XX Fungal consensus phytase protein mutant Q50T.

XX Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 KW feed additive; variant; mutein; feed; pharmaceutical.

XX Fungi.

XX Synthetic.

XX PN EP897985-A2.

XX PD 24-FEB-1999.

XX PF 15-JUL-1998; 98EP-0113176.

XX PR 24-JUL-1997; 97EP-0112688.

XX PA (HOFF) HOFFMANN LA ROCHE AG F.

XX Lehmann M;

XX WPI; 1999-134647/12.

XX Preparation of a consensus protein, especially a phytase - using
 PT programs to compare evolutionary similarity of sequences

XX Claim 9; Page -; 30pp; English.

XX This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which
 CC hydrolyses phytase to valuable feed additives, with a fully defined amino
 CC acid sequence given in the specification, or variant or mutein. The
 CC method is useful for improving protein properties by altering their
 CC sequence. The consensus protein and mutein are useful in food, feed or
 CC pharmaceutical compositions. This sequence is a mutant phytase protein of
 CC the invention which does not appear in the specification but has been
 CC created from the protein represented in AAW93380.

XX Sequence 467 AA;

Query Match 93.8%; Score 2314; DB 20; Length 467;
 Best Local Similarity 93.6%; Pred. No. 5,9e-222;
 Matches 437; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

QY 1 MGVPVLLSIATLFGSTGALGRGNHSCDTPVGGYQCFPEISHLWGTYSPPFFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshscdtpvggycfpeishlwtgtysspyfslade 60

QY 61 SAISPDPVKGRVTFVQVLSRHGARYPTSSASKAYSALIEATQKATAFKGYAELKTYN 120
 Db 61 saispdpvdkrvtfvqvlsrhgaryptsskaysalieatqknatafkgyafkty 120

QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgengmvsngikfryrykalarikivpfirasgsdrviasaeekfiqfqs 180

QY 181 KLADPCANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpcahpqaspvlnviipegagynntldhglctafeslgeddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDTVARTSDATQSPCEDLTFHDEWIQDYQLQSLG 300
 Db 241 arleahlpgvnltdedvynlmdcpfdtvartsdatspcedlthdewiqdyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpagvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNTMVSIFPAGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntmvisiffalglyngtkplsttsvesieetdgyaswtvpfagarayvemmqc 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrdrdfveglsfarsgnwaecefa 467

RESULT 15

AAW93380
 ID AAW93380 standard; Protein; 467 AA.
 AC AAW93380;
 DT 11-JUN-1999 (first entry)
 DE Fungal phytase protein consensus DNA.
 KW Phytase; consensus; myo-inositol hexakisphosphate; hydrolysis; food;
 OS feed additive; variant; mutein; feed; pharmaceutical.
 OS Fungi.
 OS Synthetic.
 PN EP997985-A2.
 PD 24-FEB-1999.
 PF 15-JUL-1998; 98EP-0113176.
 PR 24-JUL-1997; 97EP-0112688.
 PR (HOFF) HOFFMANN LA ROCHE AG F.
 PA Lehmann M;
 PI WPI; 1999-134647/12.
 DR N-PSDB; AAX23022.
 XX Preparation of a consensus protein, especially a phytase - using
 XX programs to compare evolutionary similarity of sequences
 PS Claim 8; Fig 2; 30pp; English.

XX This invention describes a novel process for the preparation of a
 CC consensus protein. The specific example given in the specification is
 CC that of a fungal phytase (myo-inositol hexakisphosphate) which hydrolyses
 CC phytase to valuable feed additives, with a fully defined amino acid
 CC sequence given in the specification, or variant or their sequence. The
 CC useful for improving protein properties by altering their sequence. The
 CC consensus protein and mutein are useful in food, feed or pharmaceutical
 CC compositions. This sequence represents the consensus phytase protein used
 CC in the method of the invention.
 XX
 SQ Sequence 467 AA;

Query Match 93.5%; Score 2308; DB 20; Length 467;
 Best Local Similarity 93.4%; Pred. No. 2.3e-221;
 Matches 436; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGVPVLLSIATLFGSTGALGRGNHSCDTPVGGYQCFPEISHLWGTYSPPFFSLADE 60
 Db 1 mgvfvllsialtfgstgaltgrgnshscdtpvggycfpeishlwtgtysspyfslade 60

QY 61 SAISPDPVKGRVTFVQVLSRHGARYPTSSASKAYSALIEATQKATAFKGYAELKTYN 120
 Db 61 saispdpvdkrvtfvqvlsrhgaryptsskaysalieatqknatafkgyafkty 120

QY 121 YTLGADDLTPFGQOMVNSGKIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
 Db 121 ytlgaddltpfgengmvsngikfryrykalarikivpfirasgsdrviasaeekfiqfqs 180

QY 181 KLADPCANPHQASPVINVIIPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
 Db 181 kladpcahpqaspvlnviipegagynntldhglctafeslgeddveanftalfapair 240

QY 241 ARLEAHLPGVNLTDDEVNLMDCPFDTVARTSDATQSPCEDLTFHDEWIQDYQLQSLG 300
 Db 241 arleahlpgvnltdedvynlmdcpfdtvartsdatspcedlthdewiqdyqlslg 300

QY 301 KYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
 Db 301 kygygagnplgpagvgfaneliarltrspvqdhstnhtldsnpatfplnatlyadfs 360

QY 361 HDNTMVSIFPAGLYNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCEAKEP 420
 Db 361 hdntmvisiffalglyngtkplsttsvesieetdgyaswtvpfagarayvemmqc 420

QY 421 LVRVLNDRVPLHCGVDKLGRCRDRDFVEGLSFARSGNWECEFA 467
 Db 421 lvrvlndrvplhgcavdklgrckrdrdfveglsfarsgnwaecefa 467

Search completed: October 26, 2001, 16:40:05
 Job time: 4956 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2001, 16:40:47 ; Search time 32.57 Seconds
(without alignments)
295.232 Million cell updates/sec

Title: US-09-488-265-31
Perfect score: 2468
Sequence: 1 MGFFVLLSIATLFGSTGT.....DFVEGLSFARSGNNECF 467

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 197339.seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2182	88.4	441	4	US-09-121-425-1
2	2170	87.9	467	4	US-09-121-425-2
3	1852	75.0	467	1	US-07-923-724-8
4	1852	75.0	467	2	US-08-609-426A-8
5	1852	75.0	467	2	US-08-374-652C-2
6	1848	74.9	467	1	US-08-151-574-32
7	1848	74.9	467	1	US-08-146-424-20
8	1848	74.9	467	1	US-08-693-709-2
9	1848	74.9	467	2	US-08-419-448-32
10	1848	74.9	467	2	US-08-819-825-3
11	1848	74.9	467	4	US-09-163-642-3
12	1841	74.6	467	4	US-09-155-855-3
13	1799	72.7	443	4	US-09-155-855-1
14	1794	72.7	443	4	US-09-155-855-2
15	1381	56.0	475	2	US-08-819-825-2
16	1381	56.0	475	4	US-09-163-642-2
17	845.5	34.3	443	3	US-08-993-359-30
18	833.5	33.8	443	3	US-08-993-359-22
19	830.5	33.7	439	3	US-08-993-359-24
20	830.5	33.7	439	3	US-09-221-654-2
21	830.5	33.7	439	3	US-08-989-358A-2
22	825	33.4	442	3	US-08-993-359-28
23	814	33.0	442	3	US-08-993-359-26
24	359	14.5	468	1	US-07-627-539G-2
25	354.5	14.4	446	1	US-07-627-539G-7
26	354.5	14.4	479	1	US-07-923-724-2
27	354.5	14.4	479	2	US-08-609-426A-2

ALIGNMENTS

RESULT 1

US-09-121-425-1
; Sequence 1, Application US/09121425
; Patent No. 6153418

GENERAL INFORMATION:

; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: Consensus Phytases 13239
; CURRENT APPLICATION NUMBER: US/09121,425
; CURRENT FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 441
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence

US-09-121-425-1

Query Match 88.4%; Score 2182; DB 4; Length 441;
Best Local Similarity 93.0%; Pred. No. 5.7e-217;
Matches 410; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

QY	27	NSHSCDTVDGGYQCFPELSHLWGTYSPFSLADESAISPDVFKGCRVTFVQVLSRHGARY	86
DB	1	NSHSCDTVDGGYQCFPELSHLWGTYSPFSLADESAISPDVFKGCRVTFVQVLSRHGARY	60
QY	87	PTSSASKAYSALIEAIQKATAFKGYAFLKTYNTLGGADLTPEGQKVNISGKRYR	146
DB	61	PTSSASKAYSALIEAIQKATAFKGYAFLKTYNTLGGADLTPEGQKVNISGKRYR	120
QY	147	YKALARKTVPIRAGSGDRVIAIAEKFTGFSQAKLADPGANPHQASFIINVIPEGAGY	206
DB	121	YKALARKTVPIRAGSGDRVIAIAEKFTGFSQAKLADPGANPHQASFIINVIPEGAGY	180
QY	207	NNTLDHGCTAFESSELDGDDVEANFTAVFAPPIARLEAHLPGVNLNDEVDVNMDCMPF	266
DB	181	NNTLDHGCTAFESSELDGDDVEANFTAVFAPPIARLEAHLPGVNLNDEVDVNMDCMPF	240
QY	267	DTVARTSDATQSPFCDLFTHDEWIQYDIQSLGKYYGYGAGNPLGPAQGVFNELIAR	326
DB	241	ETVARTSDATQSPFCDLFTHDEWIQYDIQSLGKYYGYGAGNPLGPAQGVFNELIAR	300
QY	327	LTHSPVDHSTNTILDSNPATFPLNATLADFSHDNTWYSIFFALGLYNGTKPLSTTSV	386

Db 301 LTRSPVDHTSTNHTLSDNPNATPLNATLVADFSDHNSMISIFALGLYNGTAPLSTTSV 360
Qy 387 ESIEETDGYASATVPFAARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCR 446
Db 361 ESIEETDGYASATVPFCARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGCAVDKLGRCR 420
Qy 447 DDFVEGLSFARSGGNWEECPA 467
Db 421 DDFVEGLSFARSGGNWAECEFA 441
RESULT 2
US-09-121-425-2
; Sequence 2, Application 'us/09121425'
; Patent No. 6153418
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Consensus Phytases
; FILE REFERENCE: consensus phytases 13239
; CURRENT APPLICATION NUMBER: US/09/121.425
; EARLIER FILING DATE: 1998-07-23
; EARLIER APPLICATION NUMBER: EPO 97112688.3
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:consensus
; OTHER INFORMATION: sequence
US-09-121-425-2
Query Match 87.9%; Score 2170; DB 4; Length 467;
Best Local Similarity 85.4%; Pred. No. 1.1e-215;
Matches 416; Conservative 11; Mismatches 20; Indels 40; Gaps 2;
Qy 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDVTGQYQCFPEISHLWGTYSPPFSLADE 60
Db 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDVTGQYQCFPEISHLWGTYSPPFSLADE 60
Qy 61 SAISPDVPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPDPCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 100
Qy 121 YTLGADDLTPFGBOOMVNSGIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 180
Db 101 YTLGADDLTPFGBOOMVNSGIFRYRYKALARKIVPFIRASGSDRVIASAEKFIQFQSA 160
Qy 181 KLADPGANPHQASPVIN-----VTIPEGAGYNNVLDHGLCTAFEE 220
Db 161 KLADPGSQHQASPVIDLIEAIQKNATAFKGYAFLKVIIEPGSGYNNVLDHGLCTAFED 220
Qy 221 SELGDDVEANFTAFVAPPTRARLEAHLPGVNTDDEVNLMDCPFDVTARTSDATQLSP 280
Db 221 SELGDDVEANFTALFAPARLEADPLGVTTDDEVVYLMDCPFEIVARTSDATELSP 280
Qy 281 FCDLFTHDEWIOYDYLQSLGKYGYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNH 340
Db 281 FCALFTHDEWROYDYLQSLGKYGYGAGNPLGPAQGVGFANELIARLTHSPVQDHTSTNH 340
Qy 341 TLDSPNATPLNATLVADFSDHNTWYSIFALGLYNGTAPLSTTSVESIEETDGYASMT 400
Db 341 TLDSPNATEPLNATLVADFSDHNSMTSIFALGLYNGTAPLSTTSVESIEETDGYASMT 400
Qy 401 VPFAARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGVVDKLGRCRDDFVEGLSFARSGG 460
Db 401 VPFGARAYVEMMOCEAEKEPLVRVLVNDVRVPLHGGCAVDKLGRCRDDFVEGLSFARSGG 460
Qy 461 NWEECEFA 467
Db 461 NWEECEFA 467

Db 461 NNAECFA 467
RESULT 3
US-07-923-724-8
; Sequence 8, Application US/07923724
; Patent No. 5780292
; GENERAL INFORMATION:
; APPLICANT: Nevalainen, Helena K.M.
; APPLICANT: Paloneimo, Marja T.
; APPLICANT: Miettinen-Oinonen, Arja S.K.
; APPLICANT: Torkkeli, Tuula K.
; APPLICANT: Cantrell, Michael
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Rambosek, John A.
; APPLICANT: Turunen, Marja K.
; APPLICANT: Fagerstr m, Richard B.
; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
; NUMBER OF INVENTION: in Trichoderma
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,724
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/496,155
; FILING DATE: 19-MAR-1990
; PRIOR APPLICATION DATA: US 07/044, 077
; APPLICATION NUMBER: UK 8610600
; FILING DATE: 30-APR-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1050.0240004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-923-724-8
Query Match 75.0%; Score 1852; DB 1; Length 467;
Best Local Similarity 74.3%; Pred. No. 7.9e-183;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;
Qy 1 MGVFVLLSIATLFGSTGTALGPRGNSHSCDVTGQYQCFPEISHLWGTYSPPFSLADE 60
Db 1 MGVSALLPLULLAGVTSGLVAPASRNQSCDVTGQYQCFSETSHLWGTYSPPFSLANE 60
Qy 61 SAISPDVPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SAISPDVPAGCRVTFQAQVLSRHGARYPTESKGYKYSALIEIQONVTFDCKYAFKTYN 120

QY 121 YTLGADDLTPFGQOQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAERFIEGQSA 180
 Db 121 YSLGADDLTPFGQOQELVNSGKIFRYQYRESLIRNIPFIASGSSRVIASGERFIEGFST 180
 QY 181 KLADPCANPHOASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 Db 181 KLKDPRAQPGQSPKIDVVISSASSNNTLDPTCTVFESELDATVEANFTATFAPSIR 240
 QY 241 ARLEAHLPGVNLTDDEVNLMDCPDTVARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
 Db 241 QRLNDLSGVTLTDREVTYLMDCSFDLSTVDTKLSPPCDLTHDEWIHYDYLQSLK 300
 QY 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSNHTLDSNPATPLNATLYADFS 360
 Db 301 KYGHGAGNPLGPTQGVYANELIARLTHSPVHDDTSSNHTLDSNPATPLNATLYADFS 360
 QY 361 HNTMVSIFFFALGLYNGTKPLSTTSVESTEEDTGYSASWTVPFAARAVEMMOCAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTVENITQDGFSSAWTVPFASRLYVEMMOCAEQEP 420
 QY 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECA 467
 Db 421 LVRVLNDRVPLHGGCPIDALGRCTRDSFVRLGSLFARSGGDWAECSA 467

RESULT 4

US-08-609-426A-8
 ; Sequence 8, Application US/08609426A
 ; Patent No. 5830733
 ; GENERAL INFORMATION:
 ; APPLICANT: Nevalainen, Helena K.M.
 ; APPLICANT: Paloheimo, Marja T.
 ; APPLICANT: Miettinen-Oinonen, Arja S.K.
 ; APPLICANT: Torkkeli, Tuula K.
 ; APPLICANT: Cantrell, Michael
 ; APPLICANT: Piddington, Christopher S.
 ; APPLICANT: Rambosek, John A.
 ; APPLICANT: Turunen, Marja K.
 ; APPLICANT: Fagerstr m, Richard B.
 ; APPLICANT: Houston, Christine S.
 ; TITLE OF INVENTION: Production of Phytase Degrading Enzymes
 ; TITLE OF INVENTION: in Trichoderma
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1100 New York Avenue, Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/609,426A
 ; FILING DATE: 01-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/923,724
 ; FILING DATE: 31-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/496,155
 ; FILING DATE: 19-MAR-1990
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/044,077
 ; FILING DATE: 29-APR-1987
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: UK 8610600
 ; FILING DATE: 30-APR-1986
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Reed, Grant E.

; REGISTRATION NUMBER: P-41,264
 ; REFERENCE/DOCKET NUMBER: 1050.0080001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-609-426A-8

Query Match 75.0%; Score 1852; DB 2; Length 467;
 Best Local Similarity 74.3%; Pred. No. 7.9e-183;

Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVFVLLSTATLGSTSGTALPGRNHSCDVIDGQCFPEISHLWGTYSFFSLADE 60
 Db 1 MGVSAVLLPLLYLAGVTSGLAVPASRNQSTCDTVDGQYCFSETSHLWGYAPFSLANE 60
 QY 61 SAISPDVPKGCRTVFQVLSRHGARYPTSSASKAYSALIEAQKNATAPKGYAFKLTYN 120
 Db 61 SAISPDVPAGCRVTFQAQLSRHGARYPTESKGYKYSALIEEQNVTTFDGYAFKLTYN 120
 QY 121 YTLGADDLTPFGQOQMVNSGKIFRYRYKALARKIVPFIASGSDRVIASAERFIEGQSA 180
 Db 121 YSLGADDLTPFGQOQELVNSGKIFRYQYRESLIRNIPFIASGSSRVIASGERFIEGFST 180
 QY 181 KLADPCANPHOASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPP 240
 Db 181 KLKDPRAQPGQSPKIDVVISSASSNNTLDPTCTVFESELDATVEANFTATFAPSIR 240
 QY 241 ARLEAHLPGVNLTDDEVNLMDCPDTVARTSDATQSLSPCDLTHDEWTDYDYLQSLG 300
 Db 241 QRLNDLSGVTLTDREVTYLMDCSFDLSTVDTKLSPPCDLTHDEWIHYDYLQSLK 300
 QY 301 KYGAGNPLGPAQGVGVNELIARLTHSPVQDHTSNHTLDSNPATPLNATLYADFS 360
 Db 301 KYGHGAGNPLGPTQGVYANELIARLTHSPVHDDTSSNHTLDSNPATPLNATLYADFS 360
 QY 361 HNTMVSIFFFALGLYNGTKPLSTTSVESTEEDTGYSASWTVPFAARAVEMMOCAEKEP 420
 Db 361 HDNGIISILFALGLYNGTKPLSTTVENITQDGFSSAWTVPFASRLYVEMMOCAEQEP 420
 QY 421 LVRVLNDRVPLHGGVDKLGCRKRDDEVEGLSFARSGGNWEECA 467
 Db 421 LVRVLNDRVPLHGGCPIDALGRCTRDSFVRLGSLFARSGGDWAECSA 467

RESULT 5

US-08-374-652C-2
 ; Sequence 2, Application US/08374652C
 ; Patent No. 5834286
 ; GENERAL INFORMATION:
 ; APPLICANT: NEVALAINEN, HELENA K.M.
 ; APPLICANT: PALOHEIMO, MARJA T.
 ; APPLICANT: FAGERSTROM, RICHARD B.
 ; APPLICANT: MIETTINEN-OINONEN, ARJA S.
 ; APPLICANT: TURUNEN, MARJA K.
 ; APPLICANT: RAMBOSEK, JOHN A.
 ; APPLICANT: PIDDINGTON, CHRISTOPHER S.
 ; APPLICANT: HOUSTON, CHRISTINE S.
 ; APPLICANT: CANTRELL, MICHAEL A.
 ; TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
 ; TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
 ; TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
 ; CITY: WASHINGTON

```

STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07058
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-374-652C-2

Query Match 75.0%; Score 1852; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 7,9e-183;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

1 MGVFVLLISLTLFGSTSGTALGPRGNHSCDTPVDGGYOCFPEISHLWGTYPFFSLADE 60
  ||| ||| : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 MGVSALLPLYLAGVTSLGAVPASRNQSTCDVDGQYOCFSETSHLWQYAPFFESLANE 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 SAISPDVPKGRVTFVQVLSRHGARPITSASKAYASALIEATQKNATFGKYAFILKTYN 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 SAISPDVPAGCRVTFQAQVLSRHGARYPTESKGGKYSALIEETQNNVTFDGGYAFILKTYN 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 YTLGADDLTPFGEOQWNSGKIFRYRYKALKARIVFFIRASGSDRVIASAEKFIIEGFQSA 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 YSLGADDLTPFGEGELVNSGKIFQRYESTLRNIIFFIRSSGSRVVASGENFIEGFQST 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 KLADPGANPHQASPVNVIIEPGAGYNNTLDHGLCTAFEESELGDDVEANFTAVTAPPIR 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 KLKDPRAQPGGSSPKIDWVISEASSNNTLDPGTCTVFEDSELADTVEANFTATFAPSIR 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 ARLEAHLPGVNLTDEDVNVLMDCPFQVARTSDATQLSPCDFTHDEWIOVDYLOSUG 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 QRLNENDLSGVLTIDTEVLYLMDMCSFDITSTVDTKLSPCDFTHDEWIHYLQSLK 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 KYGYGAGNPLGPAQGVGVNELIARLTHSPVODHTSTNHTLDSNPATPLNATLYADFS 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 KYIGHGAGNPLGPTQGVYANELLARLTHSPVHDDTSSNHTLDSNPATPLNSTLYADFS 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 HDNTWVSIFFALGLYNGTKPLSTTSVESIEETDGYSAWTPVFAARAYVEMMQCEAKEP 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 HDNGIISILFALGLYNGTKPLSTTTVENITQTDGFSASWTVPFASRLYVEMMQCEAKEP 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 LVRLVLNDVRVYPLHCGGVDKLGRCKRDDVFEGLSFARSGGNWEECEFA 467
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 LVRLVLNDVRVYPLHCGCPIDALGRCRTSRFVRLGSLFARSGGDWAECSA 467
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 301 KYCYGAGNPLGPAQGVGVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYCYGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMWSIFFALGNGTKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAEKEP 420
 Db 361 HDNGIISILFALGNGTKPLSTTTVENITQDGFSSANTVPPFASRLYVEMMOCEAEQEP 420
 QY 421 LVRVLNDRVPLHGGGVKGRKDDFVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGCPVDALGRCTRDSFVRGLSFARSGGDMWAECEFA 467
 RESULT 7
 US-08-146-424-20
 ; Sequence 20, Application US/08146424
 ; Patent No. 5593963
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN COIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE IN PLANTS
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/146,424
 ; FILING DATE: 02-NOV-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KENNEDY, BILL
 ; REGISTRATION NUMBER: 33,407
 ; REFERENCE/DOCKET NUMBER: 44615-20011.24
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-146-424-20
 Query Match 74.9%; Score 1848; DB 1; Length 467;
 Best Local Similarity 74.3%; Pred. No. 2e-182;
 Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;
 QY 1 MGFWVLLSIATLFGSTSGTALGPRGNHSCDVTDDGGYQCFPEISHLWGTYSPFFSLADE 60
 Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSSCDTDDGGYQCFSETHLWGQYAPFESLANE 60
 QY 61 SAISPDVPKGRVTFVQVLSRHGARYPTSSKAYSALIEIAIQKNATAFKGKYAPFLKTYN 120
 Db 61 SVISPEVPAGCRVTFQAQVLSRHGARYPTDSKGRKYSALIEIQKNATTFDCKYAPFLKTYN 120
 QY 121 YTLGADDLTPGEQOMNSGIKEYRYRALKARKIVPPFIRASGSDRVIASAEKFTGQSA 180

Db 121 YSLGADDLTPGEOELVNSGIKEYRYESLTRNIVPTIRSGSSRVIASGKKFTGQST 180
 QY 181 KLADPGANPHQASVINVIIPGAGYNNITLDHGLCTAFEESELGDDVEANFTAYFAPPIR 240
 Db 181 KLADPRAQPGQSSPKIDVIVISEASSNNTLDPGTCTVFEESELADTVEANFTATFVPSIR 240
 QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQLSPCDLTFHDEWIQDYDLSLG 300
 Db 241 QRLNDLSGVTLTDTEVTLMDMCSFDTISTSTVDTKLSPCDLTFHDEWINYDYLQSLK 300
 QY 301 KYCYGAGNPLGPAQGVGVFNELIARLTHSPVDHTSTNHTLDSNPATFPLNATLYADES 360
 Db 301 KYCYGAGNPLGPTQGVGYANELIARLTHSPVHDDTSSNHTLDSNPATFPLNATLYADES 360
 QY 361 HDNTMWSIFFALGNGTKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCEAEKEP 420
 Db 361 HDNGIISILFALGNGTKPLSTTTVENITQDGFSSANTVPPFASRLYVEMMOCEAEQEP 420
 QY 421 LVRVLNDRVPLHGGGVKGRKDDFVEGLSFARSGGNWEECEFA 467
 Db 421 LVRVLNDRVPLHGGCPVDALGRCTRDSFVRGLSFARSGGDMWAECEFA 467
 RESULT 8
 US-08-693-709-2
 ; Sequence 2, Application US/08693709
 ; Patent No. 5770413
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN COIJEN, ALBERT J. J.
 ; APPLICANT: RIETVELD, KRIJN
 ; APPLICANT: HOEKEMA, ANDREAS
 ; APPLICANT: PEN, JAN
 ; APPLICANT: SIJMONS, PETER C.
 ; APPLICANT: VERWOERD, TEUNIS C.
 ; TITLE OF INVENTION: THE EXPRESSION OF PHYTASE
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/693,709
 ; FILING DATE: 07-AUG-1996
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/146,424
 ; FILING DATE: 02-NOV-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Murashige, Kate H
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 24615-20011.10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-813-5600
 ; TELEFAX: 415-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 467 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal

```
;
;
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1-23
; OTHER INFORMATION:
;
US-08-693-709-2

Query Match          74.9%; Score 1848; DB 1; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVEVLLSIAIFLFGSTGCTALGPRGNHSCDVTGCGYQCFPEISHLWCTYSPFESLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNOSSCDTVDGQYQCFSETSHLWGOYAPFFSLANE 60

QY 61 SATSPDPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFQAQLSRHGARYPTDSKCKYSALIEEQQNTATFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEOQMVNSGKIFRYQESLTRNIVPFIASGSDRVIASGKKFIEGFQST 180

QY 181 KLADPGANPHQASPVNIVLPEAGYNNLTLDHGLCTAFESSELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVNIVLPEAGYNNLTLDHGLCTAFESSELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSLPFCDLFTHDEWIDYQLQSLG 300
Db 241 ORLENDLSGVTLTDEVTVYLMDCSFDTISTSTVDTKLSPFCDLFTHDEWIDYQLQSLK 300

QY 301 KYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNTWVSIFLALGXNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAEQEP 420
Db 361 HDNGIISILFALGXNGTKPLSTTIVENTITQDGFSSAWTVFPASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDVFEGLSFARSGGNWEECEFA 467
Db 421 LVRVLVNDRVVPLHGCGVDALGRCTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 9
US-08-419-448-32
; Sequence 32, Application US/08419448
; Patent No. 5863533
;
; GENERAL INFORMATION:
; APPLICANT: Robert F.M. Van Gortcom
; APPLICANT: Willem van Hartingsveldt
; APPLICANT: Petrus A. Van Paridon
; APPLICANT: Annemarie E. Veenstra
; APPLICANT: Rudolf G.M. Luttin
; APPLICANT: Gerardus Selten
; TITLE OF INVENTION: Cloning and Expression of Microbial
; TITLE OF INVENTION: Phytase
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1888
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,448
```

```
;
;
; FILING DATE: 10-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 24615-20026.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-419-448-32

Query Match          74.9%; Score 1848; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVEVLLSIAIFLFGSTGCTALGPRGNHSCDVTGCGYQCFPEISHLWCTYSPFESLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNOSSCDTVDGQYQCFSETSHLWGOYAPFFSLANE 60

QY 61 SATSPDPKGCRTVFOVLSRHGARYPTSSASKAYSALIEAIQKNATAFKGYAFLKTYN 120
Db 61 SVISPEVPAGCRVTFQAQLSRHGARYPTDSKCKYSALIEEQQNTATFDCKYAFKLTYN 120

QY 121 YTLGADDLTPFGEOQMVNSGKIFRYKALARKIVPFIASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEOQMVNSGKIFRYQESLTRNIVPFIASGSDRVIASGKKFIEGFQST 180

QY 181 KLADPGANPHQASPVNIVLPEAGYNNLTLDHGLCTAFESSELGDDVEANFTAVFAPPIR 240
Db 181 KLADPGANPHQASPVNIVLPEAGYNNLTLDHGLCTAFESSELGDDVEANFTAVFAPPIR 240

QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSLPFCDLFTHDEWIDYQLQSLG 300
Db 241 ORLENDLSGVTLTDEVTVYLMDCSFDTISTSTVDTKLSPFCDLFTHDEWIDYQLQSLK 300

QY 301 KYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360
Db 301 KYGAGNPLGPAQGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADFS 360

QY 361 HDNTWVSIFLALGXNGTKPLSTTSVESIEETDGYASWTVPFAARAYVEMMQCAEQEP 420
Db 361 HDNGIISILFALGXNGTKPLSTTIVENTITQDGFSSAWTVFPASRLYVEMMQCAEQEP 420

QY 421 LVRVLVNDRVVPLHGCGVDKLGRCRDRDVFEGLSFARSGGNWEECEFA 467
Db 421 LVRVLVNDRVVPLHGCPVDALGRCTRDSFVRGLSFARSGGDWAECEFA 467

RESULT 10
US-08-819-825-3
; Sequence 3, Application US/08819825
; Patent No. 5866118
;
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58661180 No. 58661180 disk of No. 5866118th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
;
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-819-825-3

Query Match 74.9%; Score 1848; DB 2; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFEISHLWGTYSPFFSLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFEISHLWGTYSPFFSLADE 60
QY 61 SAISPDPKGRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFLLKTYN 120
Db 61 SVISPEVPAGCRVTAQVLSRHGARYPTDSKGYKYSALIEEQONATTFDGYAFLLKTYN 120
QY 121 YTLGADDLTPFGEQELVNSGKIFRYRYKALARKIYVPIRASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEQELVNSGKIFRYRYKALARKIYVPIRASGSDRVIASAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVVISASNSNNTLDPGCTCTVEDESLADTVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSPFCDFLTHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVYLLMDMCSFDITSTVDTKLSPECDLTFHDEWIQDYLSLG 300
QY 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
QY 361 HDNTMVSIFALGLYNGKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCAEKEP 420
Db 361 HDNGIISILFALGLYNGKPLSTTTVENITQDGFSSAWTVPPFASRLYVEMMOCAEKEP 420
QY 421 LVRVLVNDVVPLHGGCPVDALGRCTRDTSFVRGLSFARSGGDWAECEFA 467
Db 421 LVRVLVNDVVPLHGGCPVDALGRCTRDTSFVRGLSFARSGGDWAECEFA 467

RESULT 11
US-09-163-642-3
; Sequence 3, Application US/09163642
; Patent No. 6221644
; GENERAL INFORMATION:
; APPLICANT: Berk, Randy M.
; APPLICANT: Ray, Michael W.
; APPLICANT: Klotz, Alan V.
; TITLE OF INVENTION: Polypeptides Having Phytase Activity
; TITLE OF INVENTION: And Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: No. 6221644o No. 6221644disk of No. 6221644th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/163,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,825
; FILING DATE: 18-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4758.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 867 0123
; TELEFAX: 212 867 0298
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 467 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-163-642-3

```

```

Query Match 74.9%; Score 1848; DB 4; Length 467;
Best Local Similarity 74.3%; Pred. No. 2e-182;
Matches 347; Conservative 42; Mismatches 78; Indels 0; Gaps 0;

QY 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFEISHLWGTYSPFFSLADE 60
Db 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVTGQYQCFEISHLWGTYSPFFSLADE 60
QY 61 SAISPDPKGRVTFVQVLSRHGARYPTSSASKAYSALIEATOKNATAFKGYAFLLKTYN 120
Db 61 SVISPEVPAGCRVTAQVLSRHGARYPTDSKGYKYSALIEEQONATTFDGYAFLLKTYN 120
QY 121 YTLGADDLTPFGEQELVNSGKIFRYRYKALARKIYVPIRASGSDRVIASAEKFIQFQSA 180
Db 121 YSLGADDLTPFGEQELVNSGKIFRYRYKALARKIYVPIRASGSDRVIASAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNNTLDHGLCTAFEESELGDDVEANFTAVFAPPIR 240
Db 181 KLKDPRAQPGQSSPKIDVVISASNSNNTLDPGCTCTVEDESLADTVEANFTAVFAPPIR 240
QY 241 ARLEAHLPGVNLTDDEVVNLMDMCPDFTVARTSDATQSPFCDFLTHDEWIQDYLSLG 300
Db 241 QRLNDLSGVTLTDEVYLLMDMCSFDITSTVDTKLSPECDLTFHDEWIQDYLSLG 300
QY 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
Db 301 KYGYGAGNPLGPGAGVGFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
QY 361 HDNTMVSIFALGLYNGKPLSTTSVESIEETDGYASWTVPPFAARAYVEMMOCAEKEP 420
Db 361 HDNGIISILFALGLYNGKPLSTTTVENITQDGFSSAWTVPPFASRLYVEMMOCAEKEP 420
QY 421 LVRVLVNDVVPLHGGCPVDALGRCTRDTSFVRGLSFARSGGDWAECEFA 467
Db 421 LVRVLVNDVVPLHGGCPVDALGRCTRDTSFVRGLSFARSGGDWAECEFA 467

RESULT 12
US-09-155-855-3

```


Sat Oct 27 15:25:53 2001

Sequence 3, Application US/09155855
Patent No. 6139902
GENERAL INFORMATION:
APPLICANT: KONDO, Hidemasa
APPLICANT: ANAZAWA, Hideharu
APPLICANT: KANEKO, Syunichi
APPLICANT: KANEKO, Syunichi
APPLICANT: NAGASHIMA, Tadashi
APPLICANT: TANGE, Tatsuya
TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
FILE REFERENCE: 81356/124
CURRENT APPLICATION NUMBER: US/09/155,855
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
EARLIER FILING DATE: 1997-04-04
EARLIER APPLICATION NUMBER: JP 084314
EARLIER FILING DATE: 1996-04-05
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 467
TYPE: PRT
ORGANISM: Aspergillus niger
US-09-155-855-3

Query Match 74.6%; Score 1841; DB 4; Length 467;

Best Local Similarity 73.2%; Pred. No. 1.1e-181; Mismatches 49; Indels 0; Gaps 0;

Matches 342; Conservative 49; Mismatches 76; Indels 0; Gaps 0;

QY 1 MGCVLLSLATLFGSTGTALGRNHSCHSDVDGQYCFPEISHLWGTYSPPFSLADE 60
DB 1 MGVSALLPLLYLLSGVTSGLAVPASRNQSCDVIDGQYCFSETSHLWGTYSPPFSLANK 60
QY 61 SAISPDVPGKCRVTVVOVLSRHGARYPTSSASKAYASALIEIAOKNATATKGYAFKLTYN 120
DB 61 SAISPDVPGKCRVTVVOVLSRHGARYPTSSASKAYASALIEIAOKNATATKGYAFKLTYN 120
QY 121 YTLGADDLTPFGEOMVNSGKIKFYRRYKALARKIVPPIRASGSDRVIASAEKFIQFQSA 180
DB 121 YTLGADDLTPFGEOMVNSGKIKFYRRYKALARKIVPPIRASGSDRVIASAEKFIQFQSA 180
QY 121 YSLGADDLTPFGEOMVNSGKIKFYRRYKALARKIVPPIRASGSDRVIASAEKFIQFQSA 180
DB 121 YSLGADDLTPFGEOMVNSGKIKFYRRYKALARKIVPPIRASGSDRVIASAEKFIQFQSA 180
QY 181 KLADPGANPHQASPVINVIPEGAGYNLTDLHGLCTAFESLGDVVEANFTAVFAPPR 240
DB 181 KLADPGANPHQASPVINVIPEGAGYNLTDLHGLCTAFESLGDVVEANFTAVFAPPR 240
QY 181 KLKDPRAQPGQSSPKIDWVISEASTNLTDPGCTVFEDSELADIEANFTATFVPSIR 240
DB 181 KLKDPRAQPGQSSPKIDWVISEASTNLTDPGCTVFEDSELADIEANFTATFVPSIR 240
QY 241 ARLEAHLPGVNLTDVNNLMDMCPDPTVARTSDATQSPFCDLFTHDEWIOVDYLSLG 300
DB 241 ARLEAHLPGVNLTDVNNLMDMCPDPTVARTSDATQSPFCDLFTHDEWIOVDYLSLG 300
QY 301 KYGAGNPLGPAQGVNFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
DB 301 KYGAGNPLGPAQGVNFVNELIARLTHSPVQDHTSTNHTLDSNPATFPLNATLYADF 360
QY 361 HDNTMVSIFPAGLYNGTKPLSTVSISBETDGYASWTVFAARAYVEMMOCEAKEP 420
DB 361 HDNTMVSIFPAGLYNGTKPLSTVSISBETDGYASWTVFAARAYVEMMOCEAKEP 420
QY 421 LVRVLVNDVVPVPHGCGVDKLGCRKDDFVGLSFGSGNWECEFA 467
DB 421 LVRVLVNDVVPVPHGCGVDKLGCRKDDFVGLSFGSGNWECEFA 467

RESULT 13

US-09-155-855-1
Sequence 1, Application US/09155855
Patent No. 6139902
GENERAL INFORMATION:
APPLICANT: KONDO, Hidemasa
APPLICANT: ANAZAWA, Hideharu
APPLICANT: KANEKO, Syunichi
APPLICANT: KANEKO, Syunichi
APPLICANT: NAGASHIMA, Tadashi
APPLICANT: TANGE, Tatsuya
TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE

FILE REFERENCE: 81356/124
CURRENT APPLICATION NUMBER: US/09/155,855
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
EARLIER FILING DATE: 1997-04-04
EARLIER APPLICATION NUMBER: JP 084314
EARLIER FILING DATE: 1996-04-05
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 443
TYPE: PRT
ORGANISM: Aspergillus niger
US-09-155-855-1

Query Match 72.9%; Score 1799; DB 4; Length 443;

Best Local Similarity 74.8%; Pred. No. 2.2e-177; Mismatches 330; Conservative 47; Mismatches 64; Indels 0; Gaps 0;

QY 27 NSHSCDVTVDGQYCFPEISHLWGTYSPPFSLADESAISPDVPGKCRVTVVOVLSRHGARY 86
DB 3 NQSTCDVTVDGQYCFPEISHLWGTYSPPFSLANKSAISPDVPGKCRVTVVOVLSRHGARY 62
QY 87 PTSSASKAYSALIEIAOKNATATKGYAFKLTYNVITLGADDLTPFGEOMVNSGKIKFYRR 146
DB 63 PTSSASKAYSALIEIAOKNATATKGYAFKLTYNVITLGADDLTPFGEOMVNSGKIKFYRR 122
QY 147 YKALARKIVPPIRASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPEGAGY 206
DB 123 YESLTRNVIPPIRASGSDRVIASAEKFIQFQSAKLADPGANPHQASPVINVIPEGAGY 182
QY 207 NNTLDHGLCTAFESLGDVVEANFTAVFAPPRARLEAHLPGVNLTDVNNLMDMCP 266
DB 193 NNTLDHGLCTAFESLGDVVEANFTAVFAPPRARLEAHLPGVNLTDVNNLMDMCP 242
QY 267 DTVARTSDATQSPFCDLFTHDEWIOVDYLSLGKYYGAGNPLGPAQGVNFVNELIAR 326
DB 243 DTISTSTVTDTKLSPFCDLFTHDEWIOVDYLSLGKYYGAGNPLGPAQGVNFVNELIAR 302
QY 327 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPAGLYNGTKPLSTTSV 386
DB 303 LTHSPVQDHTSTNHTLDSNPATFPLNATLYADFSDNTMVSIFPAGLYNGTKPLSTTSV 362
QY 387 ESIEETDGYASWTVFAARAYVEMMOCEAKEPILVRVLVNDVVPVPHGCGVDKLGCRK 446
DB 363 ENITQDGFSSAWTVFAARAYVEMMOCEAKEPILVRVLVNDVVPVPHGCGVDKLGCRK 422
QY 447 DDFVGLSFGSGNWECEFA 467
DB 423 DDFVGLSFGSGNWECEFA 443

RESULT 14

US-09-155-855-2
Sequence 2, Application US/09155855
Patent No. 6139902
GENERAL INFORMATION:
APPLICANT: KONDO, Hidemasa
APPLICANT: ANAZAWA, Hideharu
APPLICANT: KANEKO, Syunichi
APPLICANT: KANEKO, Syunichi
APPLICANT: NAGASHIMA, Tadashi
APPLICANT: TANGE, Tatsuya
TITLE OF INVENTION: NOVEL PHYTASE AND GENE ENCODING SAID PHYTASE
FILE REFERENCE: 81356/124
CURRENT APPLICATION NUMBER: US/09/155,855
CURRENT FILING DATE: 1998-10-05
EARLIER APPLICATION NUMBER: WO PCT/JP97/01175
EARLIER FILING DATE: 1997-04-04
EARLIER APPLICATION NUMBER: JP 084314
EARLIER FILING DATE: 1996-04-05
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 2

; LENGTH: 443

; TYPE: PRT

; ORGANISM: Aspergillus niger

; FEATURE:

; OTHER INFORMATION: Xaa at positions 2 & 3 may be any amino acid or unknown

US-09-155-855-2

Query Match

Best Local Similarity 72.7%; Score 1794; DB 4; Length 443;

Matches 329; Conservative 47; Mismatches 62; Indels 0; Gaps 0;

QY 30 SCDVGGYQCFPEISHLWCTYSPFTSLADESALSIPDKRCRVTFVQVLSRHGARYPTS 89

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 TCDVGGYQCFSETSHLWGOYAPFSLANKSALSIPDPAGCHVTFAQVLSRHGARYPTD 65

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 90 SASKAYSALIEATOKNATAPKGYAFELKTYNYTLGADDLTPFGGQOQVNSGKIFRYRYKA 149

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 66 SKGKYSALIEETQOQNTTFEGKYAEFLKTYNSLGADDLTPFGGEQELVNSGVKRYQRYES 125

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 150 LARKIYVFFRAGSDRVIASAEKFIQFQSAKLADPGAPHQASPVNVIIPGAGYNNY 209

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 126 LTRNIVFFIRSSGSRVIAAGNKFIEGFQSTKLKDPRAQFGQSPKIDVYVISEASTSNNT 185

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 210 LHGLCTAFEESELGDDVEANTAVFAPPTRARLEAHLPGVNLTDDEVDVNLMDQCPDVT 269

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 186 LDPGCTVFEDSELADDEANTATVPVIRQLENDLSGVSLTDTEVYLMDCSFDTI 245

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 270 ARTSDATQSLSPFCDLTHDEWIOYDYLQSLGYGYGAGNPLGPAQGVFNELIARLTH 329

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 246 STVDTKLSPFCDLTHDEWINDYDYLQSLKYYGHGAGNPLGPTQGVYANELIARLTH 305

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 330 SPVQDHTSNHLDSPATFELNATLYADFSDNTMVSIFALGLYNGTKPLSTTSVEST 389

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 306 SPVHDTSNHLDSPATFELNATLYADFSDNTMVSIFALGLYNGTKPLSTTSVEST 365

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 390 EETDGYASWTVPFAARAYVEMMQCEBAEKEPLRVLVNDRVPLHCGGVYDKLGRCKRDOF 449

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 366 TQTDGESSAWTVPEASRMVEMMQCEQSEPLRVLVNDRVPLHCGPVDALGRCTRDSF 425

QY 450 VEGLSFARSGGNWECFA. 467

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 426 VKGLSFARSGGDMGECFA. 443

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

RESULT 15

US-08-819-825-2

; Sequence 2, Application US/08819825

; Patent No. 5866118

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; APPLICANT: Ray, Michael W.

; APPLICANT: Klotz, Alan V.

; TITLE OF INVENTION: Polypeptides Having Phytase Activity

; TITLE OF INVENTION: And Nucleic Acids Encoding Same

; NUMBER OF SEQUENCES: 5;

; CORRESPONDENCE ADDRESSES:

; ADDRESS: No. 58661180 No. 5866118disk of No. 5866118th America, Inc.

; STREET: 405 Lexington Avenue, Suite 6400

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/819,825

; FILING DATE: 18-MAR-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 4758.200-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 867 0123

; TELEFAX: 212 867 0298

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 475 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; US-08-819-825-2

Query Match

Best Local Similarity 56.0%; Score 1381; DB 2; Length 475;

Matches 281; Conservative 61; Mismatches 107; Indels 38; Gaps 11;

QY 1 MGVF-VVLLSIATLFGSTGTALGP--RGNHSCDTVDGGYQCFPEISHLWGTYSPEFSL 57

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 6 LGSFLVLLQLQFSALL--TASPAIPFPWRKKHPND-----IARHWGQTSPEFSL 52

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 58 ADSAISPDVPGKCRVTFVQVLSRHGARYPTSSASKAYSALIEAIOKNATAFKGYAFKL 117

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 53 AEVSEISPAVPKGCRCRVEFVQVLSRHGARYPTAHKSEVYAEQLLQIQTATEFKGDAFLR 112

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 118 TYNITLGADDLTPFGGQOQVNSGKIFRYRYKALARKIVPPIRAGSDRVIASAEKFTIEGF 177

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 113 DAYHIGAONLTFRFGQEMMESGRQFYHRYREQAREITVPEVRAAGSARVIASAEFTNRGF 172

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 178 QSAKLADPGAPHQASPVNVIIPGAGYNNTLDHGLCTAFEESELGDDVE--ANFTAVEA 236

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 173 QDAKDRDRSNKDOAEFVNLIISEETGSNNLTLGLTCPAEEAP--DPTQPAEFLOVFG 230

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 237 PPIRARLEAHLPGVNLTDDEVNLMDCPFDVTARTSDAT----QLSPFCDLTHDEWIOQ 292

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 231 PRVLKXITKHPGVNLTLDEVPLFMDLCPFDVTG--SDPVLFPRLQSLSPFCHLTADDWA 288

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 293 YDYLQSLGKYGYGAGNPLGPAQGVFNELIARLT-HSPVQDHTSTNHTLSDNPATFPL 351

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 289 YDYYTLDRKYIYSHGGSAFGPSRGVGFVNELIARMTGNLFPVKDHTTVNHTLDDNPETPL 348

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 352 NATLYADFSDNTMVSIFALGLYNGTKPLSTTSVE--STEETDGYASWTVPFAARAYV 409

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 349 DAVLYADFSDNTMTGIFSAMGLYNGTKPLSTSKIQPTTGAAADGYAASWTVPFAARAYV 408

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 410 EMQOC-----EAEKEPLRVLVNDRVPLHCGGVYDKLGRCKRDRDFVEGLSFARSG 459

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 409 ELLRCETETSEEEEGEDEPFVRLVNDVPLHGCRCVDRWRCRRDEWIKGLTFAEQG 468

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 460 GNWEECF 466

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

QY 469 GHWDRCF 475

Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

Search completed: October 26, 2001, 16:40:47

Job time: 4958 sec

